

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 07:51:03 ; Search time 237.022 Seconds

(without alignments)  
335.715 Million cell updates/sec

Title: US-09-645-835a-2

Perfect score: 4376  
Sequence: 1 VKRTGYGVISVAAILATHI.....YKNGBELVYDIKTIQINP 825

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 362588 seqs, 96450795 residues

Number of hits satisfying chosen parameters: 362588

Maximum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications-AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB\_PEP:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB\_PEP:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB\_PEP:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB\_PEP:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB\_PEP:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB\_PEP:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB\_PEP:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB\_PEP:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB\_PEP:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB\_PEP:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB\_PEP:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB\_PEP:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB\_PEP:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB\_PEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4166.5	95.2	793	9	US-09-252-088-15 Sequence 15, Appl
2	3769.5	86.1	715	9	US-09-252-088-16 Sequence 16, Appl
3	1054.5	24.1	838	9	US-09-884-465A-8 Sequence 8, Appl
4	1033.5	23.6	840	9	US-09-884-465A-7 Sequence 7, Appl
5	1032	21.7	763	10	US-09-765-272-66 Sequence 66, Appl
6	949	21.7	796	10	US-09-765-272-56 Sequence 56, Appl
7	920	21.0	1039	9	US-09-884-465A-6 Sequence 6, Appl
8	864	19.7	840	9	US-09-884-465A-10 Sequence 10, Appl
9	839	19.2	447	10	US-09-765-272-182 Sequence 182, Appl
10	451	10.3	1238	9	US-09-884-465A-381 Sequence 381, Appl
11	406.5	9.3	1365	9	US-09-884-465A-382 Sequence 382, Appl
12	390.5	8.9	1378	9	US-09-884-465A-378 Sequence 378, Appl
13	359.5	8.2	999	9	US-09-884-465A-377 Sequence 377, Appl
14	359	8.2	1152	9	US-09-884-465A-379 Sequence 379, Appl
15	341	7.8	1126	9	US-09-884-465A-383 Sequence 383, Appl
16	340	7.8	999	9	US-09-884-465A-376 Sequence 376, Appl
17	337	7.5	1139	9	US-09-884-465A-380 Sequence 380, Appl
18	322	7.4	94	9	US-09-884-465A-11 Sequence 11, Appl
19	269	6.1	913	9	US-09-884-465A-384 Sequence 384, Appl

20	239	5.5	906	9	US-09-884-465A-370 Sequence 370, Appl
21	233	5.3	900	9	US-09-884-465A-334 Sequence 334, Appl
22	230	5.3	906	9	US-09-884-465A-332 Sequence 332, Appl
23	220	5.0	906	9	US-09-884-465A-372 Sequence 372, Appl
24	208	4.8	442	9	US-09-884-465A-12 Sequence 12, Appl
25	207.5	4.7	901	9	US-09-884-465A-943 Sequence 343, Appl
26	207	4.7	894	9	US-09-884-465A-943 Sequence 343, Appl
27	204	4.7	900	9	US-09-884-465A-335 Sequence 335, Appl
28	201.5	4.6	895	9	US-09-884-465A-337 Sequence 337, Appl
29	200.5	4.6	633	9	US-09-884-465A-350 Sequence 350, Appl
30	200.5	4.6	633	9	US-09-884-465A-354 Sequence 354, Appl
31	200.5	4.6	633	9	US-09-884-465A-361 Sequence 361, Appl
32	200.5	4.6	633	9	US-09-884-465A-362 Sequence 362, Appl
33	197	4.5	900	9	US-09-884-465A-339 Sequence 339, Appl
34	197	4.5	900	9	US-09-884-465A-341 Sequence 341, Appl
35	192.5	4.4	627	9	US-09-884-465A-363 Sequence 363, Appl
36	192	4.4	900	9	US-09-884-465A-367 Sequence 367, Appl
37	192	4.4	906	9	US-09-884-465A-369 Sequence 369, Appl
38	192	4.4	906	9	US-09-884-465A-373 Sequence 373, Appl
39	191.5	4.4	633	9	US-09-884-465A-351 Sequence 351, Appl
40	191.5	4.4	633	9	US-09-884-465A-352 Sequence 352, Appl
41	191.5	4.4	633	9	US-09-884-465A-357 Sequence 357, Appl
42	191.5	4.4	633	9	US-09-884-465A-358 Sequence 358, Appl
43	191.5	4.4	633	9	US-09-884-465A-359 Sequence 359, Appl
44	191.5	4.4	633	9	US-09-884-465A-360 Sequence 360, Appl
45	191.5	4.4	901	9	US-09-884-465A-345 Sequence 345, Appl

## ALIGNMENTS

RESULT 1  
US-09-252-088-15  
Sequence 15, Application US/09252088  
Publication No. US20030031682A1  
GENERAL INFORMATION:  
APPLICANT: BRODEUR, Bernard R.  
APPLICANT: RIOUX, Clement  
APPLICANT: BOYER, Martine  
APPLICANT: CHARLEBOIS, Isabelle  
APPLICANT: HAMEL, Jose  
APPLICANT: MARTIN, Denis  
TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS  
FILE REFERENCE: 8331-9002  
CURRENT FILING DATE: US/09/252,088  
EARLIER FILING DATE: 1999-02-18  
EARLIER APPLICATION NUMBER: US/60/075,425  
EARLIER FILING DATE: 1998-02-20  
NUMBER OF SEQ ID NOS: 44  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 15  
LENGTH: 793  
TYPE: PRT  
ORGANISM: Group B streptococcus  
US-09-252-088-15

Query Match 95.2%, Score 4166.5; DB 9: Length 793;  
Best Local Similarity 98.9%; Pred. No. 3.7e-289;  
Matches 787; Conservative 4; Mismatches 2; Indels 3; Gaps 2;

QY	30	MWSAKNDQAIYIDSKKAKAPKTKMDQISAEQISAOIYVKTDDGYVSHGDHY	89
DB	1	MLAKRNDQAIYIDSKKAKAPKTKMDQISAEQISAOIYVKTDDGYVSHGDHY	60
QY	90	HFYNGKVPYDAIISELMTDPNTRFKOSDVINEILDSGYVYKVGNYVYKPSKRNKI	149
DB	61	HFYNGKVPYDAIISELMTDPNTRFKOSDVINEILDSGYVYKVGNYVYKPSKRNKI	120
QY	150	RTKQIAQVAKGTEKAKGLAQAHLKSKEVAANVEAKRGRTYDDGYIFSPDIID	209
DB	121	RTKQIAQVAKGTEKAKGLAQAHLKSKEVAANVEAKRGRTYDDGYIFSPDIID	180
QY	210	DUGDAIVLVPHGNHNYITKKDLSPSELAQAQYWSOKGGRGAPSDYRPTPAPAGRRKA	269



Db 181 DLGDAIVLPHNHHYIIPKDLSPSELAAQAYWSOKRGARSDYRPT--PAPGRKA 238  
270 PIPDVTNPGGQHPDNGYHAPRPNDAQONKHOREDEFGKTFKELLDLHLDLK 329  
239 PIPDVTNPGGQHPDNGYHAPRPNDAQONKHOREDEFGKTFKELLDLHLDLK 298  
330 HVEDGLIFETQYIKSAFQYVPHGDHIIIPRSQSLPEMELADRYLAGQEDDSC 389  
299 HVEDGLIFETQYIKSAFQYVPHGDHIIIPRSQSLPEMELADRYLAGQEDDSC 358  
390 SDHSPSKDQVTHFLGRIRKAYGKGLDGRPYDSDAYVSKESHSVDSGVTAKGDH 449  
359 SEHSPSKDQVTHFLGRIRKAYGKGLDGRPYDSDAYVSKESHSVDSGVTAKGDH 418  
450 FHYIGFELQYELDEVANWYKAKGADLAALDQEGKEKPLFDTKKYSRKVTAKG 509  
419 FHYIGFELQYELDEVANWYKAKGADLAALDQEGKEKPLFDTKKYSRKVTAKG 478  
510 GYMPKRGKQYFYARDDLDLQIAFAOEMLKDKKHRYRDIYDTGIEPRLAVDSSLPM 569  
479 GYMPKRGKQYFYARDDLDLQIAFAOEMLKDKKHRYRDIYDTGIEPRLAVDSSLPM 538  
570 HAGNATYDGSSEFYIPIHDIHVYVSMLTRDQIATIKYVMOHEVRPDIWMSKRGHESG 629  
539 HAGNATYDGSSEFYIPIHDIHVYVSMLTRDQIATIKYVMOHEVRPDIWMSKRGHESG 598  
630 SVIPNVTPLDKRAGPMWQIISAEVOKALAEGRFATPDGYIFDPRDYLAKETFWAKDG 689  
599 SVIPNVTPLDKRAGPMWQIISAEVOKALAEGRFATPDGYIFDPRDYLAKETFWAKDG 658  
690 SFSPRADGSSLRITINKSDLSQAEMOQAOELAKKNGADTDTPDKREKQOAKSNNQ 749  
659 SFSPRADGSSLRITINKSDLSQAEMOQAOELAKKNGADTDTPDKREKQOAKSNNQ 718  
750 PSEASKEEKEESDPIFSLDYGIDRATLEBDHINQLOKANINDPKYLIPOEGVQYFNKN 809  
719 PSEASK--EKEESDPIFSLDYGIDRATLEBDHINQLOKANINDPKYLIPOEGVQYFNKN 777  
Qy 810 GELVTYIKTLQOINP 825  
Db 778 GELVTYIKTLQOINP 793

RESULT 2  
US-09-252-088-16  
; Sequence 16, Application US/09252088  
; Publication No. US20030031682A1  
; GENERAL INFORMATION:  
; APPLICANT: BRODEUR, Bernard R.  
; APPLICANT: RIOUX, Clement  
; APPLICANT: BOYER, Martine  
; APPLICANT: CHARLEBOIS, Isabelle  
; APPLICANT: HAMEL, Jose  
; APPLICANT: MARTIN, Denis  
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS  
; FILE REFERENCE: 8331-9002  
; CURRENT APPLICATION NUMBER: US/09/252,088  
; EARLIER APPLICATION NUMBER: 1999-02-18  
; EARLIER FILING DATE: US/60/075,425  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 715  
; TYPE: PRT  
; ORGANISM: group B streptococcus  
US-09-252-088-16

Query Match 86.1%; Score 3769.5; DB 9; Length 715;  
Best Local Similarity 98.9%; Pred. No. 7.1e-261;  
Matches 710; Conservative 4; Mismatches 1; Indels 3; Gaps 2;

Qy 108 MDPNTRFKOSDYINELDGYIKVNGNYYVLYKGRSKRNIRTKQOIAEOYAKSTKAK 167  
Db 1 MDPNTRFKOSDYINELDGYIKVNGNYYVLYKGRSKRNIRTKQOIAEOYAKSTKAK 60  
Qy 168 EKLQAVAHLSKEEVAANAENAKROGRYTTDDGYIISPDIIDDLGADALVPHGNHYIP 227  
Db 61 EKLQAVAHLSKEEVAANAENAKROGRYTTDDGYIISPDIIDDLGADALVPHGNHYIP 120  
Qy 228 KDLSPSELAAQAYWSOKRGARSDYRPTPAPGRKAPIDVTPNPGQHPDNG 287  
Db 121 KDLSPSELAAQAYWSOKRGARSDYRPTPAPGRKAPIDVTPNPGQHPDNG 178  
Qy 288 GYHAPRPNDAQONKHOREDEFGKTFKELLDLHLDLKTHVEEDGLIFETQYIKSN 347  
Db 179 GYHAPRPNDAQONKHOREDEFGKTFKELLDLHLDLKTHVEEDGLIFETQYIKSN 238  
Qy 348 AFGYVPHGDHIIIPRSQSLPEMELADRYLAGQEDDSCSDHSPSKDQVTHFLGH 407  
Db 239 AFGYVPHGDHIIIPRSQSLPEMELADRYLAGQEDDSCSDHSPSKDQVTHFLGH 298  
Qy 408 RIKAYGKGLDGRPYDSDAYVSKESHSVDSGVTAKGHFHYIGFELQYELDEVA 467  
Db 299 RIKAYGKGLDGRPYDSDAYVSKESHSVDSGVTAKGHFHYIGFELQYELDEVA 358  
Qy 468 NWYKAKGQADELAALDQEGKEKPLFDTKKYSRKVTAKGKYVMPKRGKQYFYARDDLD 527  
Db 359 NWYKAKGQADELAALDQEGKEKPLFDTKKYSRKVTAKGKYVMPKRGKQYFYARDDLD 418  
Qy 528 DLTOIAFAOEMLKDKKHRYRDIYDTGIEPRLAVDSSLPMHAGNATYDGSSEFYIPI 587  
Db 419 DLTOIAFAOEMLKDKKHRYRDIYDTGIEPRLAVDSSLPMHAGNATYDGSSEFYIPI 478  
Qy 588 DHIHVYVSMLTRDQIATIKYVMOHEVRPDIWMSKRGHEESGVYIPNVTPLDKRAGPMW 647  
Db 479 DHIHVYVSMLTRDQIATIKYVMOHEVRPDIWMSKRGHEESGVYIPNVTPLDKRAGPMW 538  
Qy 648 QIISAEVOKALAEGRFATPDGYIFDPRDYLAKETFWAKDGSEFSPRADGSSLRITINKS 707  
Db 539 QIISAEVOKALAEGRFATPDGYIFDPRDYLAKETFWAKDGSEFSPRADGSSLRITINKS 598  
Qy 708 DLQAEMOQAOELAKKNGADTDTPDKREKQOAKSNNQPSASKEEESDPIFDS 767  
Db 599 DLQAEMOQAOELAKKNGADTDTPDKREKQOAKSNNQPSASKEEESDPIFDS 657  
Qy 768 LPDYGIDRATLEBDHINQLOKANINDPKYLIPOEGVQYFNKNGELVYDITTLQOINP 825  
Db 658 LPDYGIDRATLEBDHINQLOKANINDPKYLIPOEGVQYFNKNGELVYDITTLQOINP 715

RESULT 3  
US-09-884-465A-8  
; Sequence 8, Application US/09884465A  
; Publication No. US2003007293A1  
; GENERAL INFORMATION:  
; APPLICANT: Shire Biochem, Inc.  
; APPLICANT: Hamel, Josee  
; APPLICANT: Brodeur, Bernard  
; APPLICANT: Martin, Denis  
; APPLICANT: Ouellet, Catherine  
; TITLE OF INVENTION: Streptococcus Antigens  
; FILE REFERENCE: 055190-0044  
; CURRENT APPLICATION NUMBER: US/09/884,465A  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: 60/212,683  
; NUMBER OF SEQ ID NOS: 384  
; SOFTWARE: PatentIn Version 3.1  
; SEQ ID NO 8  
; LENGTH: 838  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-884-465A-8



FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file 645835x252088.res made by msmith on Wed 28 May 103 13:41:32-PDT

Query sequence being compared:	US-09-645-835A-7 (1-816)
Number of sequences searched:	2
Number of scores above cutoff:	2

Results of the initial comparison of US-09-645-835A-7 (1-816) with  
File : US09252088Bmod.pep

SCORE	STDEV	N	SE	CONF	LOWER	UPPER
0	-3	1	10	21	-2	31
1	-1	52	63	73	84	94
2	0	1	1	1	1	1
3	0	1	1	1	1	1
4	0	1	1	1	1	1
5	0	1	1	1	1	1
6	0	1	1	1	1	1
7	0	1	1	1	1	1
8	0	1	1	1	1	1
9	0	1	1	1	1	1
10	0	1	1	1	1	1
11	0	1	1	1	1	1
12	0	1	1	1	1	1
13	0	1	1	1	1	1
14	0	1	1	1	1	1
15	0	1	1	1	1	1
16	0	1	1	1	1	1
17	0	1	1	1	1	1
18	0	1	1	1	1	1
19	0	1	1	1	1	1
20	0	1	1	1	1	1
21	0	1	1	1	1	1
22	0	1	1	1	1	1
23	0	1	1	1	1	1
24	0	1	1	1	1	1
25	0	1	1	1	1	1
26	0	1	1	1	1	1
27	0	1	1	1	1	1
28	0	1	1	1	1	1
29	0	1	1	1	1	1
30	0	1	1	1	1	1
31	0	1	1	1	1	1
32	0	1	1	1	1	1
33	0	1	1	1	1	1
34	0	1	1	1	1	1
35	0	1	1	1	1	1
36	0	1	1	1	1	1
37	0	1	1	1	1	1
38	0	1	1	1	1	1
39	0	1	1	1	1	1
40	0	1	1	1	1	1
41	0	1	1	1	1	1
42	0	1	1	1	1	1
43	0	1	1	1	1	1
44	0	1	1	1	1	1
45	0	1	1	1	1	1
46	0	1	1	1	1	1
47	0	1	1	1	1	1
48	0	1	1	1	1	1
49	0	1	1	1	1	1
50	0	1	1	1	1	1
51	0	1	1	1	1	1
52	0	1	1	1	1	1
53	0	1	1	1	1	1
54	0	1	1	1	1	1
55	0	1	1	1	1	1
56	0	1	1	1	1	1
57	0	1	1	1	1	1
58	0	1	1	1	1	1
59	0	1	1	1	1	1
60	0	1	1	1	1	1
61	0	1	1	1	1	1
62	0	1	1	1	1	1
63	0	1	1	1	1	1
64	0	1	1	1	1	1
65	0	1	1	1	1	1
66	0	1	1	1	1	1
67	0	1	1	1	1	1
68	0	1	1	1	1	1
69	0	1	1	1	1	1
70	0	1	1	1		

## PARAMETERS

	Unary	K-tuple
Similarity matrix	1	2
Mismatch penalty		20
Gap penalty	1.00	
size penalty	0.05	
off score	0	5000
randomization group	0	

## SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
79	79	65	21.21

Times:	CPU	Total Elapsed
00:00:00.00	00:00:00.00	00:00:00.00

Number of residues:	1508
Number of sequences searched:	2
Number of scores above cutoff:	2

The scores below are sorted by initial score. Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found

The 1st of best scores is:

Sequence Name	Description	Length	Score	Int. Opt.	Score	Sig.	Frame
1. US-09-252-088B-1	Sequence 15, Application	793	94	244	0.71	0	
2. US-09-252-088B-1	Sequence 16, Application	715	64	133	-0.71	0	

1. US-09-645-835A-7 (1-816) Sequence 15, Application US/09252088B-1

Sequence 15, Application US/09252088E  
GENERAL INFORMATION:

APPLICANT: BRODEUR, Bernard R.  
 APPLICANT: RIOUX, Clement  
 APPLICANT: BOYER, Martine  
 APPLICANT: CHARLEBOIS, Isabelle  
 APPLICANT: HAMEL, Josee  
 APPLICANT: MARTIN, Denise  
 TITLE OF INVENTION: NOVEL  
 FILE REFERENCE: 8331-9002  
 CURRENT APPLICATION NUMBER: US/09/252,088B  
 CURRENT FILING DATE: 1999-02-18  
 PRIOR APPLICATION NUMBER: US/60/075,425  
 PRIOR FILING DATE: 1998-02-20  
 NUMBER OF SEQ ID NOS: 45  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 15

TYPE: PRT  
ORGANISM: group B streptococcus

Initial Score	=	94	Optimized Score	=	244	Significance	=	0.71
Residue Identity	=	34%	Matches	=	295	Mismatches	=	423
Gaps	=	149	Conservative Substitutions				=	0

	10	X	30	40	50	60
MKINKYLVSAAALISVCSYELGYQARTVKENNRYSID--GQAQTAKENTLPDEYSKREGINAEQI						

MOE	10	20	30	40
X				

[illegible]

150	160	170	180	190	200
-----	-----	-----	-----	-----	-----

KRKNIRPKQIAEOYAKGTKEAKEKGLAQVAHLSKEEYAAVNEAKRGRTYTTDDGTFSPITLIDDLGDAYT  
 120 130 140 150 160 170 180 190 200 210 220 230 240 250 260 270 280 290 300 310 320 330 340 350 360 370 380 390 400

210 220 230 240 250 260  
VPHGDHYIIPKNEISASELAAEAFIS---GRG-NLSNR---TYRR--ONSNTSRTNWDPSVSNPGTT

190 200 210 220 230 240 250  
VPHGNHNYHLPKDLSPSELAAQAQWMSQSGRGARPSLYRPTAPGRKARPIPDYTPNPGGHQPDNG--

NTNTSNNSTNSQASQSD-----IDSEKQLYKPLSRHVESIGLVEFDDPAQITSRRTARGVAVPHGDHHE

260	270	280	290	300	310	320
-----	-----	-----	-----	-----	-----	-----

[illegible]

410 420 430 440 450 460 470









GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 07:51:03 ; Search time 237.022 Seconds

(without alignments)  
335.715 Million cell updates/sec

Title: US-09-645-835A-2

Sequence: 1 VKRTGYIGSVAAILLATHI.....YKNGELVTYDIKTLQIINP 825

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 362588 seqs, 96450795 residues

Total number of hits satisfying chosen parameters: 362588

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEM\_PUB pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCIT\_NEM\_PUB pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEM\_PUB pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEM\_PUB pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCITUS\_PUBCOMB pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEM\_PUB pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEM\_PUB pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEM\_PUB pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4166.5	95.2	793	9	US-09-252-088-15
2	3769.5	86.1	715	9	US-09-252-088-16
3	1054.5	24.1	838	9	US-09-884-465A-8
4	1033.5	23.6	840	9	US-09-884-465A-7
5	1032	23.6	763	10	US-09-765-272-66
6	949	21.7	796	10	US-09-765-272-56
7	920	21.0	1039	9	US-09-884-465A-6
8	864	19.7	840	9	US-09-884-465A-10
9	839	19.2	447	10	US-09-765-272-182
10	451	10.3	1238	9	US-09-884-465A-381
11	406.5	9.3	1365	9	US-09-884-465A-382
12	390.5	8.9	1378	9	US-09-884-465A-378
13	359.5	8.2	999	9	US-09-884-465A-377
14	359	8.2	1152	9	US-09-884-465A-379
15	341	7.8	1126	9	US-09-884-465A-383
16	340	7.8	999	9	US-09-884-465A-376
17	327	7.5	1139	9	US-09-884-465A-380
18	322	7.4	94	9	US-09-884-465A-11
19	269	6.1	913	9	US-09-884-465A-384

20	239	5.5	906	9	US-09-884-465A-370	Sequence 370, App
21	233	5.3	906	9	US-09-884-465A-334	Sequence 334, App
22	230	5.3	906	9	US-09-884-465A-332	Sequence 332, App
23	220	5.0	906	9	US-09-884-465A-372	Sequence 372, App
24	208	4.8	906	9	US-09-884-465A-12	Sequence 12, App
25	207.5	4.7	901	9	US-09-884-465A-343	Sequence 343, App
26	207	4.7	894	9	US-09-884-465A-337	Sequence 337, App
27	204	4.7	900	9	US-09-884-465A-335	Sequence 335, App
28	201.5	4.6	895	9	US-09-884-465A-347	Sequence 347, App
29	200.5	4.6	633	9	US-09-884-465A-350	Sequence 350, App
30	200.5	4.6	633	9	US-09-884-465A-354	Sequence 354, App
31	200.5	4.6	633	9	US-09-884-465A-361	Sequence 361, App
32	200.5	4.6	633	9	US-09-884-465A-362	Sequence 362, App
33	197	4.5	900	9	US-09-884-465A-339	Sequence 339, App
34	197	4.5	900	9	US-09-884-465A-341	Sequence 341, App
35	192.5	4.4	627	9	US-09-884-465A-367	Sequence 367, App
36	192	4.4	900	9	US-09-884-465A-333	Sequence 333, App
37	192	4.4	906	9	US-09-884-465A-369	Sequence 369, App
38	192	4.4	906	9	US-09-884-465A-373	Sequence 373, App
39	191.5	4.4	633	9	US-09-884-465A-351	Sequence 351, App
40	191.5	4.4	633	9	US-09-884-465A-352	Sequence 352, App
41	191.5	4.4	633	9	US-09-884-465A-357	Sequence 357, App
42	191.5	4.4	633	9	US-09-884-465A-358	Sequence 358, App
43	191.5	4.4	633	9	US-09-884-465A-359	Sequence 359, App
44	191.5	4.4	633	9	US-09-884-465A-360	Sequence 360, App
45	191.5	4.4	901	9	US-09-884-465A-345	Sequence 345, App

## ALIGNMENTS

RESULT 1  
US-09-252-088-15  
Sequence 15, Application US/09252088  
Publication No. US20030031682A1  
GENERAL INFORMATION:  
APPLICANT: BRODEUR, Bernard R.  
APPLICANT: RIGOUX, Clement  
APPLICANT: BOYER, Martine  
APPLICANT: CHARLEBOIS, Isabelle  
APPLICANT: HAMEL, Jose  
APPLICANT: MARTIN, Denis  
TITLE OR INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS  
FILE REFERENCE: 8331-9002  
CURRENT FILING DATE: 1999-02-18  
EARLIER FILING DATE: 1999-07-15  
EARLIER FILING DATE: 1998-02-20  
NUMBER OF SEQ ID NOS: 44  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 15  
LENGTH: 793  
TYPE: PRT  
ORGANISM: group B streptococcus  
US-09-252-088-15

Query Match 95.2% Score 4166.5; DB 9; Length 793;  
Best Local Similarity 98.9%; Pred. No. 3.7e-289;  
Matches 787; Conservative 4; Mismatches 2; Indels 3; Gaps 2;

QY 30 MGSATKQNOIYIDDSKGAAPKPTNTMOISAEBSISAOIYVKTITDGGYTSCHDY 89  
Db 1 MGLATKQNOIYIDDSKGAAPKPTNTMOISAEBSISAOIYVKTITDGGYTSCHDY 60  
QY 90 HFYNGKVPYDAIIESEELMTDPNRFKOSVDYINILDGVIKNGNYYVLKFGSKRKN 149  
Db 61 HFYNGKVPYDAIIESEELMTDPNRFKOSVDYINILDGVIKNGNYYVLKFGSKRKN 120  
QY 150 RTKOOIAEQVAKGTKEKKEGLAQAHLKREVAANVAKRGGRTTDDGYTSPDID 209  
Db 121 RTKOOIAEQVAKGTKEKKEGLAQAHLKREVAANVAKRGGRTTDDGYTSPDID 180  
QY 210 DLGPAVLPHGNHNYHLPKDDLSPELAAQAVYSQNGGARGSPDYRPTPAFGRKA 269

```

|||||
Db 181 DLGDAVIVPHGNHXYHYPKDKLSPSELAAQAQVWSQKGRGARPSPDYPT--PARGRRKA 238
QY 270 PIPDVTPNPGOGHPDNGCYHAPRPPNDASQNKHQRODFKGTRELLDQHLRLDLKYR 329
Db 239 PIPDVTPNPGOGHPDNGCYHAPRPPNDASQNKHQRODFKGTRELLDQHLRLDLKYR 298
QY 330 HVEEDGLFEPTQVYKSNAGFYVPHGHYHITIPRSQSLPEMELADRYLAGQTDNDG 389
Db 299 HVEEDGLFEPTQVYKSNAGFYVPHGHYHITIPRSQSLPEMELADRYLAGQTDNDG 358
QY 390 SDHSRPSDEKVTHTLGRIRKAYGKLDGKRPYDTSADAVYFSKESIHSVDSQVTAHGDH 449
Db 359 SEHSRPSDEKVTHTLGRIRKAYGKLDGKRPYDTSADAVYFSKESIHSVDSQVTAHGDH 418
QY 450 FHYIGFGELEQYELDEVANVMYAKQADELALDQEGKEKPLFDTKKVSRYTKDGK 509
Db 419 FHYIGFGELEQYELDEVANVMYAKQADELALDQEGKEKPLFDTKKVSRYTKDGK 478
QY 510 GYMPKPDGKDYFYARQDLDTQIAFAEOELMLDKKHRYDIDVTGIEPRLAVDVSSLPM 569
Db 479 GYMPKPDGKDYFYARQDLDTQIAFAEOELMLDKKHRYDIDVTGIEPRLAVDVSSLPM 538
QY 570 HAGNATYDTGSSFVYIPHIDHIVHPYSMLTRDQIATIKYVMOHPREVRDINSKPGHESSG 629
Db 539 HAGNATYDTGSSFVYIPHIDHIVHPYSMLTRDQIATIKYVMOHPREVRDINSKPGHESSG 598
QY 630 SYIPNVTPLDKRAGMNMWIIHSAEVOKALAEGRFATPDGTYIPRDLAKETVWMDG 689
Db 599 SYIPNVTPLDKRAGMNMWIIHSAEVOKALAEGRFATPDGTYIPRDLAKETVWMDG 658
QY 690 SFISIPRADSSLRITIKSPLSOAEMOQAOELLAKKNAGDATDTPKPKKOQADKSNENQ 749
Db 659 SFISIPRADSSLRITIKSPLSOAEMOQAOELLAKKNAGDATDTPKPKKOQADKSNENQ 718
QY 750 PSEASKEEKESDPTDLSLPDYGLDRATLEHINLOAKANIDPKYLLFQPEGVQFYNNK 809
Db 719 PSEASK-EEKESDPTDLSLPDYGLDRATLEHINLOAKANIDPKYLLFQPEGVQFYNNK 777
QY 810 GELVYDITKTLQOINP 825
Db 778 GELVYDITKTLQOINP 793

RESULT 2
US-09-252-088-16
; Sequence 16, Application US/09252088
; Publication No. US20030031682A1
; GENERAL INFORMATION:
; APPLICANT: BRODEUR, Bernard R.
; APPLICANT: RIOUX, Clement
; APPLICANT: BOYER, Martine
; APPLICANT: CHARLEBOIS, Isabelle
; APPLICANT: HAMEL, Jose
; APPLICANT: MARTIN, Denis
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
; FILE REFERENCE: 8331-9002
; CURRENT APPLICATION NUMBER: US/09/252,088
; CURRENT FILING DATE: 1999-02-18
; EARLIER APPLICATION NUMBER: US/60/075,425
; EARLIER FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 715
; TYPE: PRT
; ORGANISM: group B streptococcus
US-09-252-088-16

Query Match 86.1%; Score 3769.5; DB 9; Length 715;
Best Local Similarity 98.9%; Pred. No. 7,le-261;
Matches 710; Conservative 4; Mismatches 1; Indels 3; Gaps 2;

```

```

QY 108 MTDPNRFKOSDVINEILLDGYVIKVGNGYVYLLKPSKRRKNITRKQOIAEQAQKGTREK 167
Db 1 MTDPNRFKOSDVINEILLDGYVIKVGNGYVYLLKPSKRRKNITRKQOIAEQAQKGTREK 60
QY 168 EKLGAQVAHLKSEEVAVNNAKAKRGATTDDGYIFSEPTDIDLGAYLVPHGNHXYHYP 227
Db 61 EKLGAQVAHLKSEEVAVNNAKAKRGATTDDGYIFSEPTDIDLGAYLVPHGNHXYHYP 120
QY 228 KKDLSPELAAQAQVWSQKGRGARPSPDYPTPAPARGRRKAPIPDVTNPGOGHPDNG 287
Db 121 KKDLSPELAAQAQVWSQKGRGARPSPDYPT--PARGRRKAPIPDVTNPGOGHPDNG 178
QY 288 GYHAPRPPNDASQNKHQRODFKGTRELLDQHLRLDLKYRHYEEDGLFEPTQVYKSN 347
Db 179 GYHAPRPPNDASQNKHQRODFKGTRELLDQHLRLDLKYRHYEEDGLFEPTQVYKSN 238
QY 348 AFGYVPHGHYHITIPRSQSLPEMELADRYLAGQTDNDGSDHSKPSDEKVTHTLGR 407
Db 239 AFGYVPHGHYHITIPRSQSLPEMELADRYLAGQTDNDGSDHSKPSDEKVTHTLGR 298
QY 408 RIKAYGKLDGKRPYDTSADAVYFSKESIHSVDSQVTAHGDHFIYIGFGELEQYELDEVA 467
Db 299 RIKAYGKLDGKRPYDTSADAVYFSKESIHSVDSQVTAHGDHFIYIGFGELEQYELDEVA 358
QY 468 NVKAKQADELALDQEGKEKPLFDTKKVSRYTKDGKVGVMKPDGKDYFYARQDL 527
Db 359 NVKAKQADELALDQEGKEKPLFDTKKVSRYTKDGKVGVMKPDGKDYFYARQDL 418
QY 528 DLQIAFAEOELMLDKKHRYDIDVTGIEPRLAVDVSSLPMHAGNATYDTGSSFVYIPH 587
Db 419 DLQIAFAEOELMLDKKHRYDIDVTGIEPRLAVDVSSLPMHAGNATYDTGSSFVYIPH 478
QY 588 DHIVHPYSMLTRDQIATIKYVMOHPREVRDINSKPGHESSGVYIPNVTPLDKRAGMNM 647
Db 479 DHIVHPYSMLTRDQIATIKYVMOHPREVRDINSKPGHESSGVYIPNVTPLDKRAGMNM 538
QY 648 QIISAEVOKALAEGRFATPDGTYIPRDLAKETVWMDGSFSIPRADSSLRITIKS 707
Db 539 QIISAEVOKALAEGRFATPDGTYIPRDLAKETVWMDGSFSIPRADSSLRITIKS 598
QY 708 DLSOEMOQAOELLAKKNAGDATDTPKPKKOQADKSNENQPSSEASKEEKESDPTDLS 767
Db 599 DLSOEMOQAOELLAKKNAGDATDTPKPKKOQADKSNENQPSSEASK-EEKESDPTDLS 657
QY 768 LPDYGLDRATLEHINLOAKANIDPKYLLFQPEGVQFYNNKGEVYDITKTLQOINP 825
Db 658 LPDYGLDRATLEHINLOAKANIDPKYLLFQPEGVQFYNNKGEVYDITKTLQOINP 715

RESULT 3
US-09-884-465A-8
; Sequence 8, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 838
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-884-465A-8

```



QY 797 IFQP-----EGVOFYKNGELVYT 814  
Db 693 VEHNERPHSDNGFGNASDHVQ--RKNCGADT 723

RESULT 5  
US-09-765-272-66  
Sequence 66, Application US/09765272  
Patent No. US20020061545A1  
GENERAL INFORMATION:

APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland

COUNTRY: USA  
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/765,272

FILING DATE: 22-Jan-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083

FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 66:

SEQUENCE CHARACTERISTICS:

LENGTH: 763 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 66:

US-09-765-272-66

Query Match

Best Local Similarity 33.6%: Pred No. 2, 3e-65;

Matches 285; Conservative 101; Mismatches 249; Indels 212; Gaps 32;

23.6%: Score 1032; DB 10; Length 763;

22 SYOLGKHHNGSATKD--NOIAYTIDSSKGAAPKTYNTKMDQISAEGISAEOIYVKITDGG 80

Db 2 SYELGRHQAGOVYKESNRVSYIDGDAQKA--ENLTPEVSKREGINAEOXVYKITDGG 59

QY 81 YVTSHDHNFHYNGKVPYDAIISEELMTDPNTRFQSOVINELIDGYIKYKNGNYVYL 140

Db 60 YVTSHDHNFHYNGKVPYDAIISEELMTDPNTRFQSOVINELIDGYIKYKNGNYVYL 119

QY 141 KPGSKRNITKQOIAEOYAKGTEKKEKGLAQVAHLSKEEVAANVEAKRQGRYTTDDY 200

Db 120 KDAAHADNITKKEIKRQ-----KQERSNN-----HNSRAD--NAVAAARAQGRYTTDDY 168

QY 201 IFSPITLIDLDGAYLVPHGNHYITPKKDLSPSELAAAOAYWSQKQSGARPSDYRPTP 260

Db 169 IFNASDIETDGAYIVPHGDHYIIPKNELSASELAAEAAYWNGKO--GSRPSSSSSYN 226

QY 261 APAPGRKKAIPVPTNPGGHPDNGGCHPAPPRNDASQNHQRPDEFKGTFFKLLDQ 320

Db 227 A-----NPAOPRLSENHNLVTTP-----TYHON--OGENISSLLRE 260

QY 321 LHRDLRYRVEEDGLFEPTQYIKSNAGYVVPBGDHHIIPRSOLSPLEMLAD----- 376  
Db 261 LYAKPLSERHVESDGLFEPTQYIKSNAGYVVPBGDHHIIPRSOLSPLEMLAD----- 320

QY 377 RYLAGQTEDDSDGSHSKPSDKEVTHFTPLGHRIRIKAYGKIDGKPYTSDAYFSKESHS 436

Db 321 RYKSNHNVVD---SRPQSPQSTPSPSPQAPAPNPQAPNPIDEK-----LYKNAVVK 373

QY 437 VDKSGVTAKHGDHFHYIGFGELEQYELDEVANNVAKKGADELAAALDOQGEKPLFT 496

Db 374 V-----GDGY-----VFEENGVSRYIPAKDLSAETAAIDSKLAKQESTL--S 413

QY 497 KYVSRYTKGKGYKYMMPKQKDYFFARDDLDTQIAFQELMLDKKHRYRDIVDTGI 556

Db 414 HKLAKRTD-----LPSSDREFY--NKAYDL--LARIHQDLLDKGQVDFEALDNL 462

QY 557 EPRLAIVSSSLPMHAGNATYDTGSSFYIPIHDHIVHPYSMLTRDQJATIKYVQHPYEV 616

Db 463 E-RLK-DVXS-----DKVKLV-----XDLAFLAPIRHP-- 490

QY 617 PDWSPKGEHESGVIENVTPLDKRAGMPNMQIILHSAEYQKALAGRATPDGYTFDPR 676

Db 491 -----RLGKPNQAITYTDDDEIQVAKLGYTTEDEGYTFDPR 526

QY 677 DYLAKETFWKKDGSFSTPRADGSLRTINKSDLSQAEWQQAQELLKKNAGDATDTDKPK 736

Db 527 DITSDE-----GDYVYTPHMTHS--HWIKKDSISEBRAAQAAYAKKGL-----TPP 572

QY 737 EKQADKSNENQOPSEAKSEKESDDE--IDSLE----- 769

Db 573 STDQDSGNTFAKGAFAIYNRVKAKKVPPLDRMYPNQLQYVEVNGSLIIPHYDHYNIK 632

QY 770 ----DYGLDRA-----TLEDHINOLAQKANIDPKYLFQF-----EGVOFYN 807

Db 633 FEFWDEGLYEAPKGYTLEDLATV-----KIYVEHNERPHSDNGFGNASDHVQ--RN 683

QY 808 KNGELVYT 814

Db 684 KNGQADT 690

RESULT 6

US-09-765-272-56

Sequence 56, Application US/09765272

Patent No. US20020061545A1

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/765,272

FILING DATE: 22-Jan-2001

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083

FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P2

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 56:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 796 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 56:  
 US-09-765-272-56

Query Match 21.7%; Score 949; DB 10; Length 796;  
 Best Local Similarity 29.5%; Pred. No. 2.1e-59;  
 Matches 277; Conservative 100; Mismatches 214; Indels 348; Gaps 34;

22 SYQLGHHHGSATKDNQIAYIDDSKGAKAPRT-NKTMQISAEGISAEQIVAKITDGG 80  
 1 SYELGLYQARTYKENVRSYID--GKATQKTENLTPEVSKREGINAEQIVAKITDGG 57  
 81 YTSQHDHAFYNGKVPYPAIISELLMDPVRFKQSDVINELIDGYIKYNGVYVL 140  
 58 YTSQHDHAFYNGKVPYPAIISELLMDPVRFKQSDVINELIDGYIKYNGVYVL 117  
 141 KPGSKRNIRTRKQIAEQYAKGTKEAKGGLAQAVALHSEEAAYNAKRGQRTYTDGY 200  
 118 KDAAHADNVRTEKEINRQKQESHQ-REGG-----FRNDGAVAAARSQGRYTTDDGY 169  
 201 IFSPTDIDDDATYIVPHGNHYHYPKQDLSPELAAQAQWQSKQGG--ARPSDTRP 258  
 170 IFNASDIEDTDATYIVPHGHHYHYPKQDLSPELAAQAQWQSKQGG--ARPSDTRP 226  
 259 TPAPAPGRKAPIPDVTNPPGQHQPDNGYHAPPRPNDASQNKHQDEFGKTEKELL 318  
 227 QNSDNTSRNM-VPSYS-NPGTTNTSN-----NSNTNSQAQSQND-----IDSL 271  
 319 DQHLRLDLKRVNEEDGLFEPTQVYKSNAGYVYPHGDHHTIPRSQSLPELMELAD-- 376  
 272 KQLYKPLRSQRHVESDGLFDPQAQITSRTARGAVAPHGDHHTIPRSQSELEETARI 331  
 377 --RYLAGQREDD-----DQSGD----- 391  
 332 PLATKSNHNVWPSRPEQSPPTPEPSPGPAPRLKIDTNSLSVQLVRKVEGVEE 391  
 392 --HSPKPSDKE-VTHTFLGHR-----IKAY----- 412  
 392 KGISRYVFAKDLPESEYVKNLESKLSQESVSHITLAKKENAVPRQEFYDKAYNLLTEAH 451  
 413 -----GKGLD-----GKP----- 420  
 452 KALFYNNKGNDSFOALDKLLERLNDESTNKEKLVLDLLAFLAPITHTPERLQPSQIEYT 511  
 421 -----YDTSDAVFSKESIHSYDKSVYAKHGDNHYHGFGELEQYELDEVAN 468  
 512 EDEVRITQADLKYTTSDGTFIDEHDIISDEGDVYVPHGHSHWIKGSLDSKEVVAQA 571  
 469 WYKANG-----QADELAALDDEQGEK--PLFDTKKVSRKVTYKDGQVGYM 513  
 572 YTKREKGLIPSPDADYKANPTGDSAAATYNRKYGKRIPLV-----RLPYMV 618  
 514 PRDQKDYFARQDLDLTOIARFAEQELMLKDKKHRYRVDYVGTGIEPRLLADVSSLPMAQ 573  
 619 -----EH-----IVEYKVN-----GN 628  
 574 ATYDGSFVIRPHIDHIVVPSWL-----TRDQATTKYUQMOHEVPR---D 618  
 629 -----LIHFKDHYHNIKAFMDDHTYKAPNGYTTLEDLFATIKYVVEHDEPSPHSND 680  
 619 IWSKPGHEESGVIPNVTPLDKRAGMPNMQILHSAEEVQKALAEGRFATPDGYIRDPDV 678  
 681 GW-----GNASEHVLGCKKHSDPKNKFKADEPVEETPAE-----PE----- 718

679 LAKETFWKDGSSFIIPRADGSSLTITNKSQLEMOQAQELLAKKN-----AGDATDT-- 732  
 719 -----VPGYE-----TEKVEAQLEAEVLLAKYDSSLLKANATETTLA 755  
 733 -----DKREKQADK---SNEHQPSSEASKEE 757  
 756 GLRNNLTQIMDNNSIMAEKELLALLKGSNPSVSSEK 794

RESULT 7  
 US-09-884-465A-6  
 Sequence 6, Application US/0984465A  
 Publication No. US20030077293A1  
 GENERAL INFORMATION:  
 APPLICANT: Shire Biochem, Inc.  
 APPLICANT: Hamel, Josee  
 APPLICANT: Brodeur, Bernard  
 APPLICANT: Martin, Denis  
 APPLICANT: Charland, Nathalie  
 APPLICANT: Ouellet, Catherine  
 TITLE OF INVENTION: Streptococcus Antigens  
 PTE REFERENCE: 055190-0044  
 CURRENT APPLICATION NUMBER: US/09/884,465A  
 PRIOR APPLICATION NUMBER: 2001-06-20  
 PRIOR FILING DATE: 2000-06-20  
 NUMBER OF SEQ ID NOS: 384  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 6  
 LENGTH: 1039  
 TYPE: PRT  
 ORGANISM: Streptococcus pneumoniae  
 US-09-884-465A-6

Query Match 21.0%; Score 920; DB 9; Length 1039;  
 Best Local Similarity 26.3%; Pred. No. 3.6e-57;  
 Matches 262; Conservative 158; Mismatches 298; Indels 270; Gaps 31;

7 YIGSAATILHTHIGSVQLGKHHHGSATKDNQIAYIDDSKGAKAPRTNKTMDQISAEG 66  
 7 YIAAGSAIVYSLCATALAMHRSQENKDNKRVSYVSGSSQKS--ENLTPOVSOKEG 64  
 67 ISAQIVAKITDQGYVTSHGDNHYNGKVPYDAIISELLMDPVRFKQSDVINELID 126  
 65 IQAQIVAKITDQGYVTSHGDNHYNGKVPYDALFSELLMKQPNQQLDADIVNKG 124  
 127 GYIVKNGNHYVYLPKSGKRNITKQIAEQYAKGTKEAKGGLAQAVALHSEEAAYN 186  
 125 GYIIVKDGKYYVLDKAAHADNVRTKQELNKRQKQEHVKD--NEKVNNSVA----- 172  
 187 EAKRGRTTDDGYIFSPDTIIDLGAVLYVPHGNHYHYPKQDLSPELAAQAQWQSK 246  
 173 VARSQRTTNDGYVFNADIEDTGNAIYVPHGHHTIPKSLDSASELAQAHAHLA-- 220  
 247 QGRGAPSDYRTPAPAPGRKAPIPDVTNPPGQHQPDNGYHAPPRPNDASQNKHQ 306  
 231 -GKNMPSQSLSYSTASDNTQSAKSGTSKPAN----- 263  
 307 DEFQKTEKELLDLHLRLDKRVNEEDGLFEPTQVYKSNAGYVYPHGDHHTIPRSQ 366  
 264 ---KSENQSLKELVPSAQRYSSESDGLFDPAKITISRTPMGVALPHGDHHTIPYK 320  
 367 LSPLEMLADRYL-----AGQTEDDDS-----GSDHSPKS---DKVHTFLGHR 409  
 321 LSALEKKA-RMYVISGTSVSTNAKPNVEYSSGLSSNPSGLTTSKLS----- 371  
 410 KAYKGLDGKPYDTSADYVSKESIHSVDSGYTAKHGDNHYHYPKQDLSASELAQAHAHLA 468  
 372 -----SASDGYIFNPKDIVEETATAYIVRHGDHFHYIPKSNQIQGPLLP--N 417  
 469 WYKANGADELAALDDEQGEKPL-FDTKKVSRKVTYKDGQVGYMMPKDKQDYFYARQD 527  
 418 SLATPSPSLPINTGTSHEKHEEDGYGFD---ANRIIAEDSGVMSHGDNHYFPKDKL 473

```

OY 528 DLTOIAEGLMYLKDCKKHRYDI-----VDTGIPRLAVVSSLPJM 570
Db 474 TEEDIKAAQKILIEVKTSIHNGLBSLSSHEDDYBGNAKEMKDLDKITFEKTAGIMKQYGV 533
OY 571 AGNATPYDTGSFEVI-PHIDIHVPYSWMLTRDQIATI-----KYVMQHPREVPIWSK 622
Db 534 RESIVVNKEMKAIITYPHGDHINADPI-----DEHKRPVGHISHSNELFKPE--EGCAKK 586
OY 623 PGHE-EGGSVIPNTPPLDKRAGMNMQI-----HISABEYOKALBEG--REAP 668
Db 587 EGNVYVTGEELTNVNLKSTFNNOFNPTLANGOKRVSEFSFPRLEKKIDINMLVKLITP 646
OY 669 DGYIFD-----PRDVLAKETFF-----YMKSGSFSIPRA-- 696
Db 647 DGKYLEKUSGKVPGEVGNITANFELDQPYLPQGFYKFTIASKOYREVSYDGTFTVPSLA 706
OY 697 ----- 696
Db 707 YKMASQTIETFPFHAGDTYLRVNPQFAVPKGTDLALRVFDEFHGNALENNYKVGCEIKLP 766
OY 697 -----DGSLSRTINKSDL-----SOAEWOAOELLAKNAGADATDTDKP-- 735
Db 767 PKLWGTTRTAGNKKIPIYFMANALLDNOSTIYIEVPLLEKEN-----QTKRPSLPLQFKR 821
OY 736 KKKQOADSNSNNOQPSSEAKSEKEE-----SDDFISLDPDYGDLRATLEDHINOLA 786
Db 822 NKAQENSLDKEVVEEPTSEKVEKEKLESTGNSTNSTLEEVPVD-----PVOEKVAKFA 877
OY 787 OKANIDPKYLIFFEPEG--VOFNNKNGEV 813
Db 878 ESYGKLELVLENNMGITIELYLPGEVY 905

```

```

1 RESULT 8
2 US-09-884-465A-10
3 Sequence 10, Application US/09884465A
4 Publication NO. US20030077293A1
5 GENERAL INFORMATION:
6 APPLICANT: Shire Blochem, Inc.
7 APPLICANT: Hamel, Josee
8 APPLICANT: Brodeur, Bernard
9 APPLICANT: Martin, Denis
10 APPLICANT: Charland, Nathalie
11 APPLICANT: Ouellet, Catherine
12 TITLE OF INVENTION: Streptococcus Antigens
13 FILE REFERENCE: 055190-004
14 CURRENT APPLICATION NUMBER: US/09/884,465A
15 CURRENT FILING DATE: 2001-06-20
16 PRIORITY APPLICATION NUMBER: 60/212,683
17 PRIORITY FILING DATE: 2000-06-20
18 NUMBER OF SEQ ID NOS: 384
19 SOFTWARE: PatentIn version 3.1
20 SEQ ID NO 10
21 LENGTH: 840
22 TYPE: PRF
23 ORGANISM: Streptococcus pneumoniae
24 US-09-884-465A-10

```

```
Query Match:          19.7% Score 864; DB 9; Length 840;
Best Local Similarity 28.4%; Pred. No. 2.6e-53;
Matches 267; Conservative 144; Mismatches 273; Indels 256; Gaps 43.
```

Db	120	DAADADNRITKDELFINROKOEHVAK--NKVANSVA-----	VARSGRTYTDNGVY	167
QY	202	FSPRDIIDLDGDAIVYHGNHNHIIIPKKDLSPSYLAQAQYUNQOKGRGRPR--	DYRPT	259
Db	168	FNPADIIEDTGNATVYPRGHVYHIIIPRSDLSASELAAKHNLH--	GKNPQSOIYSTST	224
QY	260	PAPARGRRKAPITPVUTNPPGCGHORPDNGGYHAPRPRPDASOKNHORDEKFGTKFELD	319	
Db	225	PSPS-----LPI-----NPGTSH-----KH-----	240	
QY	320	QLHRLDKYRHVEEDGLIFERTYOVIKSNAGYUVPHGDHNIIPRSOLSELEMLADRY-	378	
Db	241	-----EEDGQFGDANRIIADSESGFVMSHGDHNHYEFFKRIJTEEOIKRAQKIH	288	
QY	379	-----LAGQREDDSGSDHSPSPDKETHTFELHRIKAKGKGDGRPYTSDA	426	
Db	289	EEVKTSHNGDLSDLSHEDRPSNKKEMKDDPKIEEKIAG-IMKQYV-----	KRES	339
QY	427	YVFSKESHSYDASKGVATKAGDHFRHY-----IGREL-EQYELDEVANWY-KAKQ-	475	
Db	340	IYVKE-----KNAIIYPHGDHINHAPRIDENKRVUGIGHSHSNTELPKREGVAKKEGK	393	
QY	476	---ADELAALD-----OEGKEKRLD---TKVS-----KRYTKDGKY	509	
Db	394	YVTGEELTNVNLKNTSTFNQNFNTLANKQKRVSFSRPELEKKGITIMLVKLTIPDGKY	453	
QY	510	GYMMPKDKDQV-----FYAROLDLTQIAFQEGELMLDKKHNYD---IVDGIER	558	
Db	454	---LEKVSCKVYEGEYGNANFELDQRLPRLQOT- KTIYASKDIPVSYDGTFFVPTSLAY	510	
QY	559	RLADVSSILPHMAGNATYDTGSSFVPH-----IDHIIH-----	591	
Db	511	KMASQTIYPRHAGDTYLRVNPQFAVPRKGTDALVRVEDFHGNAYLENNYKAGEIKLPIR	570	
QY	592	-----VPRYSULTR---DQIATIKYVMOHPEVAPRDIKSKGHESSVLPNV-	635	
Db	571	KLNDGTTTLAGNKIPIVTFMANAYLDNOST-YIYEIV-----IIEKENQDTPKSILPOFK	623	
QY	636	-----TPLKRGMPRMOIITISAAEEOK-ALAEGRFATPDGYI-----FDP-RDVLAK	681	
Db	634	RNKAQENSKLDEKVEER-----TSEKVEKEKLESETGNSTNSITLEVPYIVDYQVEAKV	678	
QY	682	---ETFWVK-----DGSFSIPRADSSLTITINKSDLSQAEMOQAOELAKKNAGDAP	711	
Db	679	FAESYGKMLENVLPNDQITIELYPSGEVIKK-IMADPTGBAPQNGENKRPSENGKYSTG	717	
QY	732	T---DKPRKQQAQDSNENQOQSEASKKEEKESEDFT-----DSLPTYGLDRATLED	780	
Db	738	TVENQPTNKKPADSLPE--APNEKPVPRKPNSTDGMNLPBGNGSGDPMILSALAEAPAVD	795	
QY	781	HINOLAQK-----ANIDPKYIOPRG-VQFYKNGNEI	813	
Db	796	PVQEKLEKFTASYGLGDS--VFINMGQITIELRLPSEVI	833	

```

1 RESULT 9
2 US-09-765-272-182
3 ; Sequence 182, Application US/09765272
4 ; Patent No. US20020061545A1
5 ;
6 ; GENERAL INFORMATION:
7 ;
8 ; APPLICANT: Choi, et. al.
9 ;
10 ; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
11 ;
12 ; NUMBER OF SEQUENCES: 452
13 ;
14 ; CORRESPONDENCE ADDRESSES:
15 ;
16 ; ADDRESSEE: Human Genome Sciences, Inc.
17 ;
18 ; STREET: 9410 Key West Avenue
19 ;
20 ; CITY: Rockville
21 ;
22 ; STATE: Maryland
23 ;
24 ; COUNTRY: USA
25 ;
26 ; ZIP: 20850
27 ;
28 ; COMPUTER READABLE FORM:
29 ;
30 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
31 ;
32 ; COMPUTER: HP Vectra 486/33
33 ;

```

```

OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/765,272
  FILING DATE: 22-Jan-2001
  CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/961,083
  FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
  NAME: Brookes, A. Anders
  REGISTRATION NUMBER: 36,373
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (301) 309-8504
  TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 182:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 447 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: linear
    MOLECULE TYPE: protein
  SEQUENCE DESCRIPTION: SEQ ID NO: 182:
US-09-765-272-182

```

```

Query Match      19.2% Score 839; DB 10; Length 447;
Best Local Similarity 39.9%; Pred. No. 6,4e-52;
Matches 178; Conservative 65; Mismatches 119; Indels 84; Gaps 10;

QY 25 LGHHMSATKNOIAIYDDSGKAKAKTKTKMTQISAEGISAEQIVVYITDGGYTS 84
DB 1 LMOHSQENKDNKRVSYVDGSSQSKS--ENLTPQVSQKGEIQAEQIVYITDGGYTS 58
QY 85 HGDHYHNGKVPYDAIISELLMTDPNVRKQSDVINEILDGYIKYNGVYVYLYKPGS 144
DB 59 HGDHYHNGKVPYDAIISELLMTDPNVRKQSDVINEILDGYIKYNGVYVYLYKPGS 118
QY 145 KKKNRITQOIAEVOYAKGTKEKKEKGLAQAVALHSEKVAANAENAKRGRTYTDGYSFP 204
DB 119 HADNVRTKDEINRQKQEHVKD--NEKVNQSNVA-----VASQGRYITNDGYVFN 166
QY 205 TDIIIDLDAYLVPHGNHYHYPKKDLSPEELAAQAAYWSQKRGAPSDYRPPAPAP 264
DB 167 ADIIEDTGNATYVPHGHHYHYPKSDLSASELAQAHAHA--GKNMQPSQLSYSTASD 223
QY 265 GRRKAPIDVTPNPGQGHQPDNGGYHAPRRPNDASQKHQDDEKFGKTFELLDOLHRL 324
DB 224 NNTQSVAKGSTSKRAN-----KSFNLSLKLKELYDS 254
QY 325 DLKYHNVEDGLIEPTQVYIKSMAGFYVPRGHHYHITPSQSLPELMLADRYL---A 380
DB 255 PSAQRYSESGLVDPRAKIISTRTPNGVAIPRGDHYHFIYPSKLSALEKIA--RMVPISGT 313
QY 381 GQTEDEDDSG-----GSDHSKPS---DKEVHTFLGHKIKAYGKGLDKRPDTSAY 427
DB 314 GSTVSTNAKPRNEVYSSLSLSSNPSSLTTSKELS-----SASDGY 353
QY 428 VESKESIHVSQGVYAKHGDHFNHY 453
DB 354 IFNPKDIVEETATAYIVRHGDHFNHY 379

RESULT 10
US-09-884-465A-381
; Sequence 381, Application US/09884465A
; Publication No. US2003007293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie

```

```

APPLICANT: Ouellet, Catherine
TITLE OF INVENTION: Streptococcus Antigens
FILE REFERENCE: 055190-0044
CURRENT APPLICATION NUMBER: US/09/884,465A
CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 60/212,683
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 384
SOFTWARE: PatentIn version 3.1
SEQ ID NO 381
LENGTH: 1238
TYPE: PT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1)..(1)
OTHER INFORMATION: Xaa = Methionine or nothing
NAME/KEY: MISC_FEATURE
LOCATION: (430)..(430)
OTHER INFORMATION: Xaa = Glycine or nothing
NAME/KEY: MISC_FEATURE
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa = Proline or nothing
US-09-884-465A-381

```

```

Query Match      10.3% Score 451; DB 9; Length 1238;
Best Local Similarity 19.7%; Pred. No. 1,4e-23;
Matches 221; Conservative 150; Mismatches 312; Indels 440; Gaps 41;

QY 77 TDQGYVTSQDHYHFNKGVYDAIISELLMT----- 109
DB 36 TARGVAAPHGNHNF-----IPYEQMSELEKRIARIIPRYRSNHWYDSDRPEPSQPTP 91
QY 110 -----DENYRKQSDVINEIL-----DGYIKYNG--NYVYLYK----- 142
DB 92 EPSPSPQAPADNPAPAPSNPIDKLVKCAVRKVGDDGYVEENGSRVYIPAKNLSAETAGI 151
QY 143 -----GSKRKN-----RTKQOIAEVOYAKGTK---EAKKEK 169
DB 152 DSKIAKQESLSHKLGAKTLLPSSDRFYKATVDLARIHQDLDN--KGRQVDFALDN 209
QY 170 GLAQAVALHSEKVAANAENE-----AKRQGRYTTDDG 199
DB 210 LLEELKDVSSDKVYIVDDIAPLPIRHPRLGKPNQAQIYTTDEIQVAKLACKYTTEDG 269
QY 200 YIPSPITIIDLDAYLVPHGNHYHYPKKDLSPEELAAQAAYWSQK----- 246
DB 270 YIPDPRTISDEGDVAYVTPMHTSHWIKKDSLSEARAAQAAYAKKEKGLTPSTDHQDSG 329
QY 247 --GGRGAPSDYRPTAPARRRKAPIDVTPNPGQGHQPDNG----- 287
DB 330 NTEAKGAEIYNVKA-----KKVPLDRPVMLOYTVEKKNOSLLIIPYDHNHNKFEW 384
QY 288 ---GYHAP-----PRPD-----A 299
DB 385 FDEELYEAPRGYTLLEDLATVKKYVEHPNERPHSDNGFGNASDHYVXXNMQPSQLSYSTPA 444
QY 300 SOKKHO-----RDEFGKTFKELLDOLHRLDLYRHNVEDGLIEPTQVYIKSMAGFY 351
DB 445 SDNNTQSVAKGSTSKRANSENLQSLKLELYDSPSAQRYSESGLVDPRAKIISTRTPNGV 504
QY 352 VPHGDHYHITPSQSLPELMLADRYL---AGQTEDEDDSG-----GSDHSKPS-- 396
DB 505 ALPHGDHYHFIYPSKLSALEKIA--RMVPISGTGSTVSTNAKPRNEVYSSLSLSSNPSSL 563
QY 397 --DKEVHTFLGHKIKAYGKGLDKRPDTSAYVESKESIHVSQGVYAKHGDHFNHY- 453
DB 564 TTSKELS-----SASDGYIFNPKDIVEETATAYIVRHGDHFNHYIP 603
QY 454 GFGELEQYELDEVAANWYAKAGADELAALDDQGEKPL--PFTKVVSKVYRKDGWVGVM 512
DB 604 KSNQIGQPTLPN--NSLATPSPSLPINPGTSHKHEEDDQGFDP---ANKITAEDESGFV 657

```

QY 513 MPKDGKDYFVARQDLDLQIAFAEOELMLDKKHRYDI-----VDTG 555  
 DB 658 MSHGDHNYFFPKDLTEBOIKAQKHLEVKYTSINGDLSLSSHHODYPGNAKKEKDLDDK 717  
 QY 556 IEPLAVDVSSLPMAHAGNATYDTGSSFVI-PRIDHIVVPYSWLTROJATI-----K 607  
 DB 718 IEKXIAGIMKQYGVKRESIYVVKREKNAIIPHGHHADPI-----DEHKPVGIGHSHN 772  
 QY 608 YVMOHPEVRPRDMSKPGHE-ESSGVIIPNVTPLDKRACMPNQI-----IHSAEV 656  
 DB 773 YELFKPE--EGVAKKEGKKVYTGEBELTNVNLKNSIFNNONFTLANGOKRVSEFPPEL 830  
 QY 657 OKALAE--FRATPDGYIFD-----PRDYLAKETP-----684  
 DB 831 EKKLGINMLVLTITPDGKVLKESGVKPGEGVGNIANFELDQPLPGOTFEYTTIASKDYP 890  
 QY 685 -VWKDGSFSIPRA-----696  
 DB 891 EYSYDGTFTVPTSLAYKMASQTIFFPHAGDTYLAVNPQFAVPKGTALVRFDEFGNA 950  
 QY 687 -----DSSSLRTIKSDL-----SQABMOQAQOELLAKKNAGD 728  
 DB 951 YLENNYKVGETIKLPIPKLNOGTTTACGKIPVTFMANAYLDNOSTYIYEVPILEKEN---1007  
 QY 729 ATDTDKP-----KEKOADKSNENQOPEASKEEKE-----SDPFIDSLPDY 771  
 DB 1008 -QIDKPSILFOFKRNKQOENSKLDEKVEEPKJSEKVEKELSETGNSSTLEEVPTV 1065  
 QY 772 GLDRATLEDHINQIAOKANIDPKYLIFOPES-VQFYNNKNGELV 813  
 DB 1066 D-----PVQEKVAKFAESYGMKLENVLFMMDGTIELYLPESGEVI 1104

RESULT 11  
 US-09-884-465A-382  
 : Sequence 382, Application US/09884465A  
 : Publication No. US20030077293A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Shire Biochem, Inc.  
 : APPLICANT: Hamel, Josee  
 : APPLICANT: Brodeur, Bernard  
 : APPLICANT: Martin, Denis  
 : APPLICANT: Charland, Nathalie  
 : APPLICANT: Ouellet, Catherine  
 : TITLE OF INVENTION: Streptococcus Antigens  
 : FILE REFERENCE: 055190-0044  
 : CURRENT APPLICATION NUMBER: US/09/884,465A  
 : CURRENT FILING DATE: 2001-06-20  
 : PRIOR APPLICATION NUMBER: 60/212,683  
 : PRIOR FILING DATE: 2000-06-20  
 : NUMBER OF SEQ ID NOS: 384  
 : SOFTWARE: PatentIn version 3.1  
 : SEQ ID NO 382  
 : LENGTH: 1365  
 : TYPE: PRT  
 : ORGANISM: Artificial Sequence  
 : FEATURE:  
 : OTHER INFORMATION: Unknown Organism  
 : NAME/KEY: MISC\_FEATURE  
 : LOCATION: (1)..(1)  
 : OTHER INFORMATION: xaa = Methionine or nothing  
 : NAME/KEY: MISC\_FEATURE  
 : LOCATION: (557)..(557)  
 : OTHER INFORMATION: xaa = Glycine or nothing  
 : NAME/KEY: MISC\_FEATURE  
 : LOCATION: (558)..(558)  
 : OTHER INFORMATION: xaa = Proline or nothing  
 : US-09-884-465A-382

Query Match 9.3%; Score 406.5; DB 9; Length 1365;  
 Best Local Similarity 17.8%; Pred. No. 2.4e-20;  
 Matches 222; Conservative 156; Mismatches 310; Indels 557; Gaps 43;

QY 77 TDQGVYTSBGDHYHEFYNGKVPYDAIISELLMT-----109  
 DB 36 TAGVAVAPRIGNHYHF-----IPYEQMSLEKRIARILIPLRNSNMWVPDSRPEESPQPTP 91  
 QY 110 -----DPNRFKQSDVINEL-----DGVYKVNQ-NYYVYLKP-----142  
 DB 92 EPSRPOPARPNQPARSPNIDELKVAERKVGDCGVEEENVSRIIPAKNLSTAETANGI 151  
 QY 143 -----GSKRKNI-----RTQOIAQVAKGT---EAREK 169  
 DB 152 DSLAKQESLSHKLGAKKTDLSSDREFFYNKAVDILARIHQDLDLN--KGRQVDEALDN 209  
 DB 210 LERLADSDSKVKLVDDILAFAPIRHBERLGRPNAQITTYTDEIQAQKAGKYTTEG 269  
 QY 200 YIFSPTDIIDLDGATLVPHGNHYHYPKKDLSPLSLAAQAYMSOK-----246  
 DB 270 YIDPRDITSDEGDAYVTPHMTHSHWIKKDSLSEARAAQAYAKEGLTPPSTHODSG 329  
 QY 247 -OGRGA-----251  
 DB 330 NTAQAGAEALYNKVAKKAVPLDRMPYNLOTYEVKNSLIIPHYNHNIKFEWFDEGL 389  
 QY 252 -----RPSDTR 257  
 DB 390 YEAPKGYTLEDLATVYVYEHENRPHSDNGFNASDHVQRNKNQADVTQTERPSEK 449  
 QY 258 P-----TP--APARGKRAPIDVTP--NPQGNQP-----284  
 DB 450 POTEKPEETPREBKQSEKPESEKPTPEEPESEPESEPOVETEKEVKELREADDLG 509  
 QY 285 -----DNG-----GYHPAPPEND 298  
 DB 510 KIDPPIIKSNAKETLNGKNNLLFGTDNNNTIWAFAEKLLALLKESKXXXNQPQSLSYSS 569  
 QY 299 ASONKHORDEFKGTKE-----ELDQHLRLDLKTRHVEDDLIEFPQVYKSNMF 349  
 DB 570 TASDNNQSVAKGSTSKPANKSENLOSLLKELYDSQAQYSSDGLVDPAKIISRPN 629  
 QY 350 GYVPHGDHYHIIIPRSQSLPELEMLADRYL--AGOTEDDD--GSDSKSPS 396  
 DB 630 GVALIPGDHYHFIPTSKLSALEKIA-RWVPISGTSYSTNAKPRBYVSSLSGSSNS 688  
 QY 397 ----DKEVTHFLGHRILKAYGKGLDGKPYDSDAYVESKESHSYVSKGVTAKHGDHYH 452  
 DB 689 SLTTSKELS-----SASDGYIFNPQDIVEETATAYIVHGDHFHY 728  
 QY 453 I-GGGELEQYELDEVANWYAKQADELAALDOOGKEKPL-FDVKKYSRKVTKGKYG 510  
 DB 729 IPKSNQIGQPTLPN--NSLATPSPSLPIINGTSHKHEEDGYGFD--ANRIIADESG 782  
 QY 511 YMPKDGKDYFVARQDLDLQIAFAEOELMLDKKHRYDI-----VD 553  
 DB 783 FVMSHGDHNYFFPKDLTEBOIKAQKHLEVKYTSINGDLSLSSHHODYPGNAKKEKDLDD 842  
 QY 554 TGIEPLAVDVSSLPMAHAGNATYDTGSSFVI-PRIDHIVVPYSWLTROJATI-----606  
 DB 843 KIEBKTAGIMKQYGVKRESIYVVKREKNAIIPHGHHADPI-----DEHKPVGIGHSH 897  
 QY 607 -KYVMOHPEVRPRDMSKPGHE-ESSGVIIPNVTPLDKRACMPNQI-----IHSAE 654  
 DB 898 SNYELFKPE--EGVAKKEGKKVYTGEBELTNVNLKNSIFNNONFTLANGOKRVSEFPPE 955  
 QY 655 EVOKALAE--FRATPDGYIFD-----PRDYLAKETP-----684  
 DB 956 ELEKLGINMLVLTITPDGKVLKESGVKPGEGVGNIANFELDQPLPGOTFEYTTIASKD 1015  
 QY 685 -VWKDGSFSIPRA-----696  
 DB 1016 YPEVSYDGTFTVPTSLAYKMASQTIFFPHAGDTYLAVNPQFAVPKGTALVRFDEFGNA 1075



QY 697 -----DSSLRITINKSDL-----SOAEWQAOELLAKKNA 726  
 Db 1076 NMLENNYKVGELIKLPKLNOSTTTRAGNKIPYTFMANAYLDNOSTYIVETPILEKEN- 1134  
 QY 727 GDAITDTPK-----KEKQADKSNENQOPSEASKEEKE-----SDDEPIDSLP 769  
 Db 1135 ----QDCKPSILPQFKRNKAQENSKLDEKVEEPTSEKVEKEKLESTGNSNSTLEVP 1190  
 QY 770 DYGLDRATLEDHINOLAOKANIDPKLIFQPEG-VQFYKNGELY 813  
 Db 1191 TVD-----PVOEKVAKFAESYGMKLENVLFNMDGTIELYLPSEVI 1231

## RESULT 12

US-09-884-465A-378  
 ; Sequence 378, Application US/09884465A  
 ; Publication No. US2003007293A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shire Biochem, Inc.  
 ; APPLICANT: Hamel, Josee  
 ; APPLICANT: Brodeur, Bernard  
 ; APPLICANT: Martin, Denis  
 ; APPLICANT: Charland, Nathalie  
 ; APPLICANT: Ouellet, Catherine  
 ; TITLE OF INVENTION: Streptococcus Antigens  
 ; FILE REFERENCE: 055190-0044  
 ; CURRENT APPLICATION NUMBER: US/09/884,465A  
 ; CURRENT FILING DATE: 2001-06-20  
 ; PRIOR APPLICATION NUMBER: 60/212,683  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 384  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 378  
 ; LENGTH: 1378  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Unknown Organism  
 ; NAME/KEY: MISC FEATURE  
 ; LOCATION: (1)..(1)  
 ; OTHER INFORMATION: Xaa = Methionine or nothing  
 ; NAME/KEY: MISC FEATURE  
 ; LOCATION: (570)..(570)  
 ; OTHER INFORMATION: Xaa = Glycine or nothing  
 ; NAME/KEY: MISC FEATURE  
 ; LOCATION: (571)..(571)  
 ; OTHER INFORMATION: Xaa = Proline or nothing  
 ; US-09-884-465A-378

Query Match 8.9%; Score 390.5; DB 9; Length 1378;

Best local Similarity 17.6%; Pred. No. 3.4e-19;

Matches 221; Conservative 149; Mismatches 318; Indels 569; Gaps 42;

QY 77 TDGQYVTHSGDHHFYNGKVPYDAIIEF----- 105  
 Db 37 TARGVAAPHGNHVF-----IPYEQMSELEKRIARIIPLRYSNHWVPDSRPEQSPQSTP 92  
 QY 106 ----LMTDNYRFRKQSDVINEL-----DGVIKNG-NNYYLKP----- 142  
 Db 93 EPPSLDPAHPNPAPSPNPIDELKLVKAVKVGDDYFEENGVSRTYIPAKDLSAETRAAGT 152  
 QY 143 ----GSKRKNI-----RTKQOIAEOVANGTK---EAKK 169  
 Db 153 DSKLAKQESLSHKLAGKKTDLPSDSREFYNKAYDALLARITHQDLDN--KGRQVDFEVLN 210  
 QY 170 GLAQAHLKKEVAANVE-----AKRGRYTTDDG 199  
 Db 211 LLEKIKVNSSKYKLVLDIAFLAPIRHPRERLGRPNQOITYTDEIQVAKLAKYTTEDG 270  
 QY 200 YIFSPDIDLDGAYLVPHGNHNYTPKKDLSPSELAAQAAVWSOK----- 246  
 Db 271 YIFPDITSDGAYVTPHMTSHWIKKQSLSEAEAAQAAVAKKEGLTPPSTHDGSG 330

QY 247 --QCGAR----- 252  
 Db 331 NTEAKGAAITNRKAKAKVPLDMRPYNLOYTEVKNKSLIIPHYDHNHKKFEMPEGL 390  
 QY 253 ---PSDY-----RPTPAPAG-----RRKAPIDVYPPNGQGHQPD 285  
 Db 391 YEAPKGSLELDLATVKKYVEHPNERPHSDNGFNASDHVKKKADQDSKDEDEKHEV 450  
 QY 286 NGCYHP----- 291  
 Db 451 SEPTPESEDEKENHAGLNPASDNLYKOSTTEETEEAEADTTDEALIPQVENSYNAKIA 510  
 QY 292 -----APRP----- 296  
 Db 511 DAELLEKVTDPSTIRQNAETLITGLKSSLLTGKDNNTISAENVSLALLAKESQAPAP1X 570  
 QY 297 -----NDASQNKQ-----RDEFKGTKEILDQJLRDLKRYHVEEDGL 337  
 Db 571 XNMPQSOLSYSTASDNNTQSAKSTSPANKSENLSLKLKELYDSPSAQRYSESGLV 630  
 QY 338 PEPTQVTKSMNAGYVPHGDHNIIPRSOLSPLEMLADRYL---AGQTEDDDS----- 388  
 Db 631 FDPAKIISRTNGVAIPGHDHIFIPYKLSALEKIA-RNVPISGTSYSTVSTAKRNEV 689  
 QY 389 ----GSDHSKPS---DKEVTHFLGHRIRKAYGKGLDGKPYDSDAYVFSKESIHSYDKS 440  
 Db 690 VSLGSLSNPSLTSTSKLS-----SASDGYTFNPKDIVEETAT 729  
 QY 441 GYTAHGHGHFHTI-GFGELOYELDEVANWYKAGQADELAAALDQEGKRPKPL-FDTRK 498  
 Db 730 AYIVRHGDHFHYIPKSNQIGOPTLPN--NSLATPSPSLIMPSTSHENHEEDGYGF-- 784  
 QY 499 VSRKTKDQGVGMPKRGKRDYFVARQDLDLQIAFAQOELMDKRYRDI----- 551  
 Db 785 -ANRIIADESEFVSHGDHNHYFFPKDLTERQIKAAQKHLEVYTSINGDLSSHOD 843  
 QY 552 -----VDTGIEPLAVDVSSLPMHAGNATYTGSSFYI-PHIDHIVVPYSLWTR 600  
 Db 844 YPGNAKEMKDLKITEEKIAGIMQYGVKRESIYVNRKKNM1IYPHGDHNDADPT----- 898  
 QY 601 DQIATL-----KYVMQHPYRPRDMSKPGHE-EGSVIAPVPTLDRACMPNQI--- 649  
 Db 899 DEHKPVGIGHSNYSLEKPE--EGVAKKEGKVTYGEELTNVWLLKNSFPNNQNTFLA 956  
 QY 650 ----HSAEVOKALAEQ--REFATPDGYTFD-----PRVLA 680  
 Db 957 NQOKRVSEFPPELEKKLQIMLVKLIITPDGKLVLEKVGKVGEGVGNIANPELDQPLP 1016  
 QY 681 KETF-----VKDGSFSIPRA----- 696  
 Db 1017 GQFKYTIASKDYPREVSYDGTFTVPTSLAYKMASOTIFYPFHAGDTYLRVNPQAVPQGT 1076  
 QY 697 -----DSSLRITINKSDL-----SOAEW 714  
 Db 1077 DALVRVEDEFHGNAYLENNYKVGELIKLPKLNOSTTTRAGNKIPYTFMANAYLDNOSTY 1136  
 QY 715 QOAOELLAKKAGDADITDKP-----KEKQADKSNENQOPSEASKEEKE----- 760  
 Db 1137 IYEVPILEKEN-----QDCKPSILPQFKRNKAQENSKLDEKVEEPTSEKVEKEKLESTG 1191  
 QY 761 ----SDDEPIDSLPDYGLDRATLEDHINOLAOKANIDPKLIFQPEG-VQFYKNGELY 813  
 Db 1192 NSTSNSTLEVPVVD-----PVOEKVAKFAESYGMKLENVLFNMDGTIELYLPSEVI 1244

## RESULT 13

US-09-884-465A-377  
 ; Sequence 377, Application US/09884465A  
 ; Publication No. US2003007293A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shire Biochem, Inc.  
 ; APPLICANT: Hamel, Josee  
 ; APPLICANT: Brodeur, Bernard

```

APPLICANT: Martin, Denis
APPLICANT: Charland, Nathalie
APPLICANT: Ouellet, Catherine
TITLE OF INVENTION: Streptococcus Antigens
FILE REFERENCE: 055190-0044
CURRENT APPLICATION NUMBER: US/09/884,465A
CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 60/212,683
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 384
SOFTWARE: PatentIn version 3.1
SEQ ID NO 377
LENGTH: 999
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (1)..(1)
OTHER INFORMATION: Xaa = Methionine or nothing
NAME/KEY: MISC_FEATURE
LOCATION: (570)..(570)
OTHER INFORMATION: Xaa = Glycine or nothing
NAME/KEY: MISC_FEATURE
LOCATION: (571)..(571)
OTHER INFORMATION: Xaa = Proline or nothing
US-09-884-465A-377

```

Query Match 8.2% Score 359.5; DB 9; Length 999;

Best Local Similarity 22.3%; Pred. No. 3.5e-17;

Matches 180; Conservative 114; Mismatches 282; Indels 231; Gaps 37;

```

QY 60 QISAEGLSAEIVVKTIDGQVTVTHGDIHFYNGKVPYDAIISSELMITDPNY----- 113
DB 260 QFAVPRGTDA---LVKVFDE---FHGNAYLENNYKV-----GEIKLPIKLNQGTTR 305
QY 114 ---RKQSDVINELIDG---YVIKV-----NGNTYVYLAKGSRKRNIRTKQOIAEOV- 159
DB 306 TAGNKIPVTFMNAVALDNOSTYIVVEPILEKENQTDKPSILFQFRKNKAQENSKLDEKVE 365
QY 160 -AKGTKEKEKGLAQVAHLK---EEVAANVEAKRO--GRYTTDGY----- 200
DB 366 EPKTSKVEKEKLESTGNTSSTLEEVPTVDPOEKAKPAFESGMLENVLFNMDGTI 425
QY 201 -IFSPF-----DIIDGDAYLVPHGNHYHYPKKDLSPELAAQAYWSQKQGRGAP 253
DB 426 ELYLPSGEVYIKKNMADFTGEA---POGN-----GENKPSENGKYSTGVENQPTENKP 475
QY 254 SDYRTPRAPAPGRKAPI-PDVTNPGQCHOPDNGGYHP-----APP----- 294
DB 476 ADULP---EAPNEK---PKAPENSTNGMLNPEGNVGSDPMLDPALEAPAVDPQEKLEK 530
QY 295 -----RPNDASQNKHQDEFEKGKTFKLLDQLHRLDKYRHV 331
DB 531 PTASVGLGDSVIFNMDGTIELRLPSGEVYIKKNLSDFIAXXDIDLLKQYKLPISQKHV 590
QY 332 EEDGLFEPTQVYIKSNAGYVVRPHGDHNIIPRSQLSPELMELAD---RYLAGOTEDDD 387
DB 591 ESDGLIFPQAQITSTRAGVAVPHGNHYHIFPEQMSLEKRIARIIPLRYRSHMWVD- 649
QY 388 SSGSDHKSQSDKEVHTFLGHRIKAYGKGDGKPRYTSDAVYFSKSHSVYKSGVTAKHG 447
DB 650 ---SREDESPQPTPEPSPSPQAPNPQAPNPIDEK---LVKAVKRV-----G 694
QY 448 DHFNHIGFGLBYELDEVANVMKAKGADLAALDOEGEKEKFLPTKRVSRKVTKDQ 507
DB 695 DCY-----VFDENGSVRYTPAKNLSAETAGIDSKLAKQESL--SHLGAKTID-- 741
QY 508 KGVMMPRDKQYFYARQDLTQIAFADQELMLKDKKHRYRDVDTGIEPRLAVDSSL 567
DB 742 ---LPSDREYF---NKAYDL--LARIHODLLDNKGRQVDFEALDNLLE--RLK-DVSS- 789
QY 568 PMHAGNATGDTGSSFVLIHIDHIVVPSWMLTRDQIATIKYVMQNPBEVRPIWSPGHEE 627

```

```

DB 790 -----DKVKLV-----DIIAFLAPIHNP----- 809
QY 628 SGSVIPIVNTPLDKRAGPMNQIISAEVYOKALAEGRFATPDGYPDPDLAKETFFVWK 687
DB 810 -----BLCKFMAQITFTYDDEIQVAKLAGKYTEDGIEFPDPRITSD----- 851
QY 688 DGSFIPRAGSSSLRTJINKSDLSQAEWQOQOELLAKKNAGDADITDCKPKKQQAADSNEN 747
DB 852 GDATVYTPMTHS--HWIKDSLSAEERAAQAVAKKEGL-----TPSTHDQSGNTE 902
QY 748 QQPSEASKEEKEESDDF-IDSLPDYGL 773
DB 903 AKGAELATYNRVYKAKKVPDLDRMP-YNL 928

```

RESULT 14

US-09-884-465A-379

Sequence 379, Application US/09884465A

Publication No. US20030077293A1

GENERAL INFORMATION:

APPLICANT: Shire Biochem, Inc.

APPLICANT: Hamel, Josée

APPLICANT: Brodeur, Bernard

APPLICANT: Martin, Denis

APPLICANT: Charland, Nathalie

APPLICANT: Ouellet, Catherine

TITLE OF INVENTION: Streptococcus Antigens

FILE REFERENCE: 055190-0044

CURRENT APPLICATION NUMBER: US/09/884,465A

CURRENT FILING DATE: 2001-06-20

PRIOR APPLICATION NUMBER: 60/212,683

PRIOR FILING DATE: 2000-06-20

NUMBER OF SEQ ID NOS: 384

SOFTWARE: PatentIn version 3.1

SEQ ID NO 379

LENGTH: 1152

TYPE: PRF

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: MISC\_FEATURE

LOCATION: (1)..(1)

OTHER INFORMATION: Xaa = Methionine or nothing

NAME/KEY: MISC\_FEATURE

LOCATION: (344)..(344)

OTHER INFORMATION: Xaa = Glycine or nothing

NAME/KEY: MISC\_FEATURE

LOCATION: (345)..(345)

OTHER INFORMATION: Xaa = Proline or nothing

US-09-884-465A-379

Query Match 8.2% Score 359; DB 9; Length 1152;

Best Local Similarity 20.7%; Pred. No. 4.7e-17;

Matches 196; Conservative 138; Mismatches 290; Indels 322; Gaps 50;

```

QY 32 SATKQNOIAYIDSKGKAKAPKTKNT-----MDQISAEGLSAEQIV-----VK 75
DB 356 STADNNNTQGV--AKGSTSKP--ANKSENLQSLKELYDSSAQRYSESQGLVDPKAIIS 412
QY 76 ITDQGYVTSQGDHYHFNPGNVKVPIDATIS--BELMTDPMNTRFKOSDVINELIDGYIVKVG 134
DB 413 RTPGVAVIRPHGDHNF---IPYSKLSALEKEI---ARWVPISG 449
QY 135 NYUYULKPGSKRKRNIRTKQOIAEOVAK--GRKEAKEKGLAQVAHLKSEVVAANVEAKRGR 193
DB 450 T-----GS---TVSTNAKPRNEVYSSLSGSSNPSSLITTSKLS----- 484
QY 194 YTTDGYIFSPDTIIDGDAYLVPHGNHYHYPKKDLSPELAAQAYWSQKQGRGAP 253
DB 485 -SASDGYIFPKPDIVETATVAVIRHGDHNYTPK-----SNQIQGPTLP 528
QY 254 SDYRTPRAPAPGRKAPIPDVTNPGQCHOPDNGGYHPAPRPNDASQNKHQDEFGKKT 313

```

```

Db      529  NNSLAPSPS-----LPI-----NPGTSH-----KH-----550
QY      314  FKELLQHLRLDKYRNVEDGLIEPTQVYIKSNAGYVVPDGHYIIPRSQLSPLEME 373
Db      551  -----BEDGIFANNTIADDESGFVMSHGDNHYEFKKDLIEEQIK 592
QY      374  LADRY-----LAGOTEDDSDGSDSKSPDKEVTFTLGHRIKAYGKGLDCKP 420
Db      593  AAOKHLEEVKTHNGLDLSSSHODYPGNAKEMKDLCKIEEKING-IMKQGV-----645
QY      421  YQSDAVYFSKESIHVSQSYTAKGHFH-----IGFGL-EOYELDEVANWY-470
Db      646  --KRESIVNKE-----KNATYPRGDHMHADPIDENKPVGIGHSHSNYELFKPEEGVA 697
QY      471  KAKQ-----ADELAALD-----QEOGKEKLPD-----TKVS-----RV 503
Db      698  KKEGNKYVGEELTNVNLKSTENONFTLANQOKRVSFSPELEKLLINMLVKLI 757
QY      504  TQDGKGYMMPKDGKDY-----FYARDQLDTQIAFAEOELMKDKKHRYD---IV 552
Db      758  TPDGKV--LEKVSQKVFGEVGNIAFELDQYLPQQT-XYTIASKDQPEVSYDGTFTV 814
QY      553  DNGIEPRLAVDVSLLPMHAGNATYDTGSSFYTPH-----IDH-----591
Db      815  PISLAKYMASQITFYPRFHAGDTYLRVNPQFAVPKGTDALVRVDFHGNAYLENNYKVOE 874
QY      592  -----VPRYSMLTR---DQIATIKYVMOHPEVRPDMWSKPGHESSGS 630
Db      875  IKLPIPKLNGOTTRRAGNKIPIYTFMANATLDNOST--YIVEY-----ILEKENDTKPS 927
QY      631  VIPNV-----TPLDKRAQMPNMQIHSABEVOK-ALAEGRATPDGYI-----FDP- 675
Db      928  ILPQRKRNKAQENSXIDEKVEEPK-----TSEKVEKEKLESTGNSTSLTEEVPTVDP 982
QY      676  RQVLAK--ETFEVWK-----DGSSEIPRADGSSLRTINKSDLSQAEMQAOELLAKN 725
Db      983  QKRVAKFAESYKMLENLVFNMDGTIELYLPSEGVYIKR-NMADFGEAPDQNGENKPS 1041
QY      726  AGDADT--DKPREQOADSNNQOPSEASKKEEKESDDF-----DSLDPDGLD 774
Db      1042  GKVSTGYVNGCTENKPADSLPE--APNEKPYKPNSTNDGMLNEGAVGSDPMLDPAL 1099
QY      775  RATLEDHINQAOQ-----ANIDKRYLIFQPEG--VOFYNNKNGELV 813
Db      1100  EAPAVDPVOEKLEKFTASYGLDLS--VIFNMDGTIELRLPSGEVI 1143

RESULT 15
US-09-884-465A-383
; Sequence 383, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 383
; LENGTH: 1126
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Unknown Organism

```

```

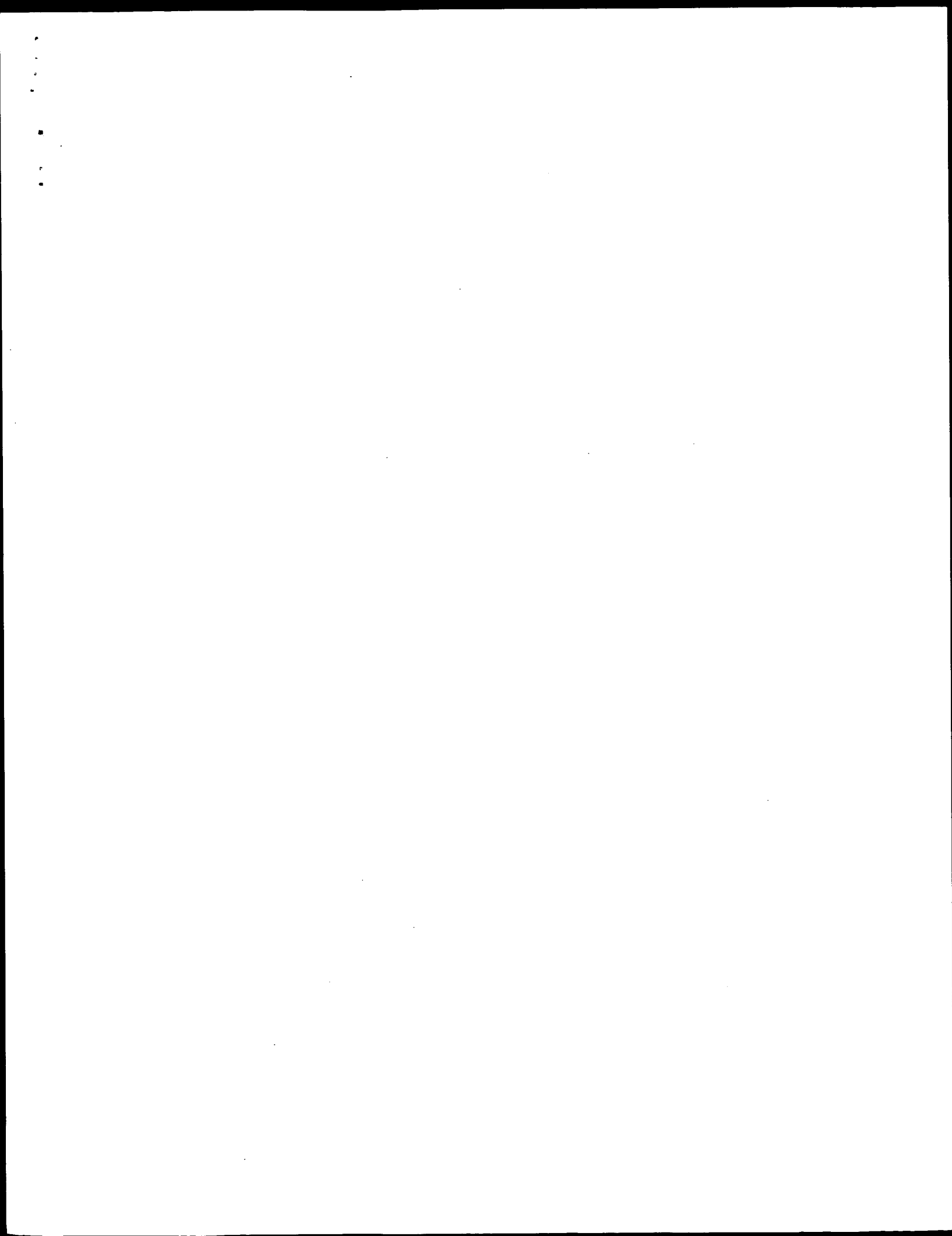
; NAME/KEY: MISC FEATURE
; LOCATION: (1) (1)
; OTHER INFORMATION: Xaa = Methionine or nothing
; NAME/KEY: MISC FEATURE
; LOCATION: (557) (557)
; OTHER INFORMATION: Xaa = Glycine or nothing
; NAME/KEY: MISC FEATURE
; LOCATION: (558) (558)
; OTHER INFORMATION: Xaa = Proline or nothing
US-09-884-465A-383

Query Match      7.8%; Score 341; DB 9; Length 1126;
Best Local Similarity 25.0%; Pred. No. 8.8e-16;
Matches 138; Conservative 66; Mismatches 177; Indels 170; Gaps 23;

QY      317  LLDLHLRLDLKRYHVEDGLIEPTQVYIKSNAGYVVPDGHYIIPRSQLSPLEME 376
Db      6  LKOLYKLPLOSQHVESDGLIFDPAQITSPARGAVAPDHGNYHFIFEQMSELEKRIAR 65
QY      377  -----RYLAGOTEDDSDGSDSKSPDKEVTFTLGHRIKAYGKGLDCKPYPDTSAYVFSKE 432
Db      66  IYPLRYBNHWPV-----SRPEESPQTPPEPSPQAPAPNPQAPSNPIDEK---LVKE 118
QY      433  SIHSYDSQGYTAKGHDFHYIGFGELEQYELDEVANWVAKAKQAOELAAALDOEGKEK 492
Db      119  AVRKV-----GDGY-----VFEENGYSRYIPAKNLSAETAAAGIDSKLAQES 160
QY      493  LPDTRKYSRKATKQKGYMMPKDKGDTFYARDQLDTQIAFAEOELMKDKKHRYDY 552
Db      161  L--SHKLGAKRTD-----LPSSDREFY--NKAYD--LARIHQDILLDKKGRQVDFEAL 207
QY      553  DTGIEPRLAVVSSLLPMHAGNATYDTGSSFYIPIHDHIVVPSWMLRQDQATIKYMOH 612
Db      208  DNLE-RLK-DVSS-----DKVKLV-----DQILAFLAIRH 237
QY      613  PEVRPDIWSKPGHESSGVIPNVTPLDKRAQMPNMQIHSABEVOKALAEGRATPDGYI 672
Db      238  PE-----RLGKPRQAQITTYDDELQVAKLAGKVTTEGCI 271
QY      673  FDPBDVLAKETFWKDGDSFSTRADGSSLRTINKSDLSQAEMQAOELLAKNAGDADT 732
Db      272  FDPDITSDS-----GDAYVTPHMTHS--HWIKKDSLSAERAAQAAYAKKGL-318
QY      733  DKPREQOADSNNQOPSEASKKEEKESDDF--IDSLP-----769
Db      319  -TPPSTHODSGNTEAGAEIAYRVAKAAKVPLDRMPYMLQYTVYEVKNSLIIPHYDH 377
QY      770  -----DYGLDRA---TLEDHINQAOQANIDPKYLIIFOP-----EGV 803
Db      378  HNIKFEWFDEGLYAPKGYTLEDILATV-----KYVEHNERPHSDNGFGNASDHY 429
QY      804  QFYNNKNGELV 814
Db      430  Q-RNKNQADT 439

```

Search completed: May 19, 2003, 08:23:52  
 Job time: 249.022 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 07:51:03 ; Search time 12.8853 Seconds

(without alignments)  
1883.841 Million cell updates/sec

Title: US-09-645-835A-2

Perfect score: 4376  
Sequence: 1 VKRTYGISVAAILLATHI.....YKNGELVYDIKTIQINP 825

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1032	23.6	763	4	US-08-961-083-66
2	949	21.7	796	4	US-08-961-083-56
3	839	19.2	447	4	US-08-961-083-182
4	143.5	3.3	1861	2	US-08-790-912-4
5	142.5	3.3	1588	5	PCT-US93-07261-11
6	142.5	3.3	1588	5	PCT-US93-07261-16
7	133	3.0	2465	2	US-08-596-291-3
8	133	3.0	2465	2	US-09-100-804-3
9	131.5	3.0	1541	4	US-08-296-791-3
10	131.5	3.0	1541	5	PCT-US95-10661A-3
11	129	2.9	893	4	US-09-514-302-4
12	129	2.9	1938	4	US-09-514-302-2
13	128.5	2.9	1076	2	US-08-867-941-19
14	128.5	2.9	1076	4	US-09-074-658-19
15	127	2.9	765	2	US-08-663-112-2
16	137	2.9	783	6	5231168-2
17	125	2.9	10182	4	US-09-134-001C-3159
18	124	2.8	706	4	US-09-059-584-46
19	124	2.8	1848	4	US-08-296-791-6
20	124	2.8	1848	5	PCT-US95-10661A-6
21	123.5	2.8	2366	1	US-08-480-604A-10
22	123.5	2.8	2366	2	US-08-405-496A-10
23	123.5	2.8	2366	4	US-08-915-136-10
24	123.5	2.8	2366	4	US-08-957-310-10
25	121.5	2.8	1702	4	US-08-296-791-5
26	121.5	2.8	1702	5	PCT-US95-10661A-5
27	121	2.8	506	3	US-09-032-365A-19

28	121	2.8	608	1	US-08-480-604A-21	Sequence 21, Appl
29	121	2.8	608	2	US-08-405-496A-21	Sequence 21, Appl
30	121	2.8	608	4	US-08-915-136-21	Sequence 21, Appl
31	121	2.8	608	4	US-08-957-310-21	Sequence 21, Appl
32	121	2.8	609	1	US-08-480-604A-30	Sequence 30, Appl
33	121	2.8	609	4	US-08-915-136-30	Sequence 30, Appl
34	121	2.8	2466	5	US-09-080-855-12	Sequence 12, Appl
35	121	2.8	2466	5	PCT-US94-09943-2	Sequence 2, Appl
36	120.5	2.8	542	1	US-08-701-380-2	Sequence 2, Appl
37	120.5	2.8	542	3	US-09-032-365A-13	Sequence 13, Appl
38	120.5	2.8	747	1	US-07-854-596B-10	Sequence 40, Appl
39	120.5	2.8	1151	4	US-09-134-001C-3242	Sequence 3242, Ap
40	120.5	2.8	1181	2	US-08-488-940-2	Sequence 2, Appl
41	120	2.7	2485	4	US-09-290-640-46	Sequence 46, Appl
42	117.5	2.7	672	3	US-09-040-843-4	Sequence 4, Appl
43	117.5	2.7	672	4	US-09-621-853-4	Sequence 4, Appl
44	117.5	2.7	823	4	US-09-134-001C-4081	Sequence 4081, Ap
45	117.5	2.7	866	3	US-09-040-843-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-08-961-083-66  
; Sequence 66, Application US/08961083  
; Patent No. 6159469  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,083  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 66:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 763 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-961-083-66  
  
Query Match 23.6%, Score 1032, DB 4; Length 763;  
Best Local Similarity 33.6%, Pred. No. 1.8e-75;  
Matches 285; Conservative 101; Mismatches 249; Indels 212; Gaps 32;  
DB 22 SYOLKHHNGSKTKD-NOIAYIDSKSKAKAPKTKMTDOIISAEIGTISAEOIVKTIIDOG 80  
2 STELRHQAQGVKSKSNVSYITDGDQAGOKA--ENLTPDEVSKREGINAEOXVTKITIDOG 59





TOPOLOGY: linear  
MOLECULE TYPE: protein  
us-08-790-912-4

Query Match 3.3%; Score 143.5; DB 2; Length 1861;  
Best Local Similarity 18.3%; Pred. No. 0.014;  
Matches 190; Conservative 166; Mismatches 328; Indels 353; Gaps 56;

```

QY 36 DNOIAYIDDSKGAAPK-----TNKTMD-QISAE-GISAEIYVITIDGQVYTHSDHY 89
DB 199 DKK---IPDNQONANVDKALNOKKLDYSFQNGKLNQVGYNTIE----- 244
QY 90 HFYNGKVPYDAIISBELMTDPNRFKOSDVINEILDGVIKVN-----GNYVYL--KP 142
DB 245 -----PQDEVLSGR--VAKPELLYKETSITETIAYGQIQENDLAEQYTRVQDEKPP 295
QY 143 GSKRRKIR-----TKQIAEOVAKGTAKREKGL-----AQVA 175
DB 296 GKKIEVVRIFTVDNAEVSHEVISTKIEETPKIVKGTKLEAPSEKPYTSMNVQPEOVA 355
QY 176 HLSK-----EEVAAVNEAKROGRYT--TDGYIFSPDTIIDDLG-----DAYLVP 218
DB 356 PLPEYTGVSQAIIVEPEOVASLPE-----YSGTISGAIIVEPEQLEPELIGVQSGAIYEP 409
QY 219 HGNHYHYIPKID-----LSPSELAAQAQVWSOKGRCARPSDYRPTP-----A 261
DB 410 E---QVPLPEYTGQAGAVVSPQVAPLPEYTGQSGAIIVEPAQVTPLEPYTGVSQGA 465
QY 262 PAPGRKAPIDP-----VTPNP-----CQ 280
DB 466 IYKPAQVPLPEYTGQSGAIIVEPEOVAPSPPEYTGQAGAIIVEPROVASLPEYTGVSQGA 525
QY 281 GHQPD-----NGYHPAPRPND-----ASQNKHQRDE--FKGTFKELLD--QLHRL--D 325
DB 526 IYEPPEVEPEPQYTGIEPAEAENPPEKAKQEPKQEPKNEKLNKNSVDELXLAD 585
QY 326 LAYR--HVEEDGL-----IEP--TOYKSNAGFYVYPHGDHYIIPRS 365
DB 586 GKYYKQVHSDAIIPSNQENFYVKKSKFKDVFPLISSIVYKTKDQPV-----YKTTASA 640
QY 366 QLSPELMELADRYLAGQTEDDSGSDHSPKDEYTH--TFIGHRIKAYGKGLDGKPYTSS 424
DB 641 E--KLKQDVNNKY-----EDNFTYFLAKKAREVTNFTSFSNLVQALNNNLNGTYTLAA 692
QY 425 DAVYFSKESIHSDVSGYTAHKGHDEHYIGFELQYELDEVANVYKAK-----GOADE 478
DB 693 -----SLNAVEVELENG-----ASSYIKGFTGKLFSGKDG 723
QY 479 LAAALDQEGCKEPLFDTRYKVS-----RKYTKDGK--VGYMMKRDGKDYIYARDQDL 529
DB 724 KNYAI--YNLKKPFLDLSAATVENTLTKDVNISGKTIDIGAL-----ANEANNA 770
QY 530 TOIAFAEOELMLDKK-----HYRYDIVD-----TG 555
DB 771 TRINNVHYDGVLAGERGIGGLVWKADNSKISMSSEKGRILVNSYETRKAPINIGLVLGOLTG 830
QY 556 I-----EPRLAVDVSSLPMHAGNATYDTG-----SSEVYIPHIDIH----- 591
DB 831 INALVDKSKATITISS-----NADSTQVGLAGLVEKDALISNSYAEGINNVKRFSGVA 887
QY 592 -VVPYSMLTRDQIATIKYMQHPEVAPRDIWSPKRGHESSSVYVNTPLDKRAGMNN--WQ 648
DB 888 GAGYILM--DRDS--SEERHAGRLHNLVLDINVNGNAISGYHYRGMRITDSYSNKDNRYK 945
QY 649 I-IHSAEEVOKALAEGRFATPDGYIPEPRDVLAKETFWKDGCSFSIPRADGSLTLTKNS 707
DB 946 VLEKEEVYTKSLER-----GTILDVSOIASKSEI-----NSLAPKYEITLLITNKES 997
QY 708 DLSQAEQMOAOELLAKKNA-----GDATDTDK--PKEQOAKDSNENQO 749
DB 998 DSKYVDYQASRALAYKNIKELLPLFYKNATIKYKGLVKEVDSTLVEKEILSAVMKMDNEV 1057
QY 750 PSEASKSEKESDFTIDSLPYGLDR--ATLEDHINQLAQKANIDPKYLIPOPEGVORFN 807

```

```

DB 1058 ITDIASHKE--AANKLLIHYKHDSSEKLDLTYQSDFSKLAIE-YVCGDTGLYTPN--QFLQ 1113
QY 808 KNGELVYTYDITLQOIN 824
DB 1114 NHHSSIVNEVLPDLKAVD 1130

```

## RESULT 5

```

PCT-US93-07261-11
Sequence 11, Application PC/TUS9307261
GENERAL INFORMATION:
TITLE OF INVENTION: PFEMP3 MALARIA ANTIGEN, ANALOGS, ANTIBODIES AND USES THEREOF
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: John H. C. Blasdale
STREET: One Giralda Farms
CITY: Madison
STATE: New Jersey
COUNTRY: USA
ZIP: 07940-1000
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.5
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07261
FILING DATE: 19930805
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/927,531
FILING DATE: 07-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Blasdale, John H. C.
REGISTRATION NUMBER: 31,895
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-822-7398
TELEFAX: 201-822-7039
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1588 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-07261-11

```

Query Match 3.3%; Score 142.5; DB 5; Length 1588;

Best Local Similarity 20.0%; Pred. No. 0.014;

Matches 174; Conservative 120; Mismatches 341; Indels 233; Gaps 38;

```

QY 27 KHHGSAATKDNQIAYIDDSKGAAPKTN-----KTMDQISAEISAEQIYVITIDGQY 81
DB 4 KEALQKQTEKKEKARNALKEKIKLEQKKNDQAQKDLTKKESQDSSEKSLKKEVNGEAL 63
QY 82 VTSHGDIHYHFNQVVPYDAIISBELMTDPNRFKOSDVINEILDG--YIAYKNGN----- 135
DB 64 KEKE-----NKETLKKLELQENKEKEENKIKINDNDALKKNGDKDD 106
QY 136 -YUYVLKPGSKRNIR-----TKQIAEOVAKGTREAKEK-----GLAQVA 175
DB 107 KVIYKPKPESEYKDLKEMELKEKEFKIÖHLKDYERKEKRRNWLNSLRDKLREIÖLE 166
QY 176 HLSKEEVAANVEAK-----RQGRYTTDDGYIFSPND-----IIDDL 211
DB 167 KLNQLESAINELKERRASRRPMVVKMQRGKDEVDWEMIKKYDDEQAEKNGTKDEELKDK 226
QY 212 GDAY--LVP--HGNHYHYIPKIDLSPSELAAQAQVWSOKGRCARPSDYRPTPAPAPGR 266
DB 227 GDGYEIVETKFGYMRNALDELID--EYERYEKKRYLKEDEGSG--DLK----- 272
QY 267 RKAPIDPVTTPRGQGHQPDNGCYHAPRPPNDASQNKHÖDEFFKGTFFKLLDQLHRLDL 326

```



Db 273 ---DVEEKLEETGCVFREKRPFTTILIVKRRKNNKQDKKLKLEKKEKLLAEEPPDEKKITKL 329  
 QY 327 KYRHHVEEGGLIFEPQVILKSN-AFGYVVPFGCDHHIIPRQSLP---EMELADRYLAG 381  
 Db 330 K---DSDKRVVVPVNNKKSSEFPDKFRAPDCKRTMEFYRLSELPFIVPRKDNELA---VCG 382  
 QY 382 QTEDDSGSDHSKSPDKRVTHTF-----LGHNIIKAYGCKL----- 416  
 Db 383 DSMD-----SKVNGKLLKSTFNPFRKRRKKLKERKKMOELHAKKKKNNKYKOKLLEREKR 435  
 QY 417 ---DCKPDTSDAYVFSKESHSVDKSGVTAKHGDHPHYTGFGELQYELDEVANVWK-- 471  
 Db 436 ENPDEPPLNTEPIHVIHPSDL-MDK-GENKSAGHPPEYOPTGLKREYESHSVKDYOLE 492  
 QY 472 -----AKGQDELALALDOQGERKLEPTTKVSRKVTYDQGVGYMMPDGCDYF 521  
 Db 493 HEPTPKLPEYERKGVHSR-EYOLDHEPPTKLPEYKGVHSRKYOLDNEVRDELLPEYEGHV 551  
 QY 522 YARDOLD---LTOIAFAEOLML-KDKKHRYDVIDVTGIEPRLA--VDVSSLPMHAGNA 574  
 Db 552 SREYOLDNEGSPSTLKEYDQTELAKGDKDITNKPHESSVDEYQFTELAKKDKDITNKP-HESVD 610  
 QY 575 TYDGGSSVVIHDIHIVPYSWLTTRDQIATIKYVQHPPEYRPD-----IWSK 622  
 Db 611 EYD-----QSELAKGDKDITNKPHESSVDEYDQTELAKGKEVTNK 648  
 QY 623 PGHEESGVIPNVTPLDKRAGMPNMQIHSASEYQKALAGREAT-----PGYIFDPR 676  
 Db 649 P-HENLEBY--NEIDLAKGKEVTNKPHESSVDEYQSELAKGDKDITNKPHESSVDEY---DQ 702  
 QY 677 DYLAKEFTFWKDGSSFSIPRADGSLRTINKSDLSQAEMQQAELL--AKKNAGDAVDTDK 734  
 Db 703 TELAK-----GKEVTNKAARENLEEYNETDLA-----GKEVTNKAARENLEEYNETDL 749  
 QY 735 PREKQADSKSNENQOPEASKEEKESED 762  
 Db 750 AKGKEVTNKAHEN-----LEEYNETD 770  
 RESULT 6  
 PCT-US93-07261-16  
 : Sequence 16, Application PC/TUS9307261  
 : GENERAL INFORMATION:  
 : TITLE OF INVENTION: PFEAP3 MALARIA ANTIGEN, ANALOGS, ANTIBODIES AND USES THEREOF  
 : NUMBER OF SEQUENCES: 23  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: John H. C. Blasdale  
 : STREET: One Giralda Farms  
 : CITY: Madison  
 : STATE: New Jersey  
 : COUNTRY: USA  
 : ZIP: 07940-1000  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: Apple Macintosh  
 : OPERATING SYSTEM: Macintosh 6.0.5  
 : SOFTWARE: Microsoft Word 5.1a  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/US93/07261  
 : FILING DATE: 19930805  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 07/927,531  
 : FILING DATE: 07-AUG-1992  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Blasdale, John H. C.  
 : REGISTRATION NUMBER: 31,895  
 : REFERENCE/DOCKET NUMBER: DX0288K  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 201-822-7398  
 : TELEFAX: 201-822-7039  
 : INFORMATION FOR SEQ. ID NO.: 16:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 1663 amino acids

```

; TYPE: amino acid
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: peptide
;
; ORIGINAL SOURCE:
;
; ORANISM: Plasmodium falciparum
;
; STRAIN: Malayan Camp
;
PCT-US93-07261-16

```

Query Match	3.38;	Score 142.5;	DB 5;	Length 1663;
Best Local Similarity	20.08;	Pred NO	0.015;	

Matches 174; Conservative 120; Mismatches 341; Indels 233; Gaps 38;

2/ KHHMGSATKDNQIAYIDDSKSGKAKAPKTN-----KTMDQISAEEGISAQIIVKITDQGY 81

4 KEALKQTEKNEKARNALKEKKLKEQKKNDQAQAKDLTKKESQDSSSEKSLKEKVNGEAL 63

82 V1SHGDHYHFNKGVPYDAIISELLMTDPNRYRFKQSDVINEILDG--YVIKVNGN----- 135

04 KEKE-----NKETLKKKELENQKEKEKEKNIKDNNDKALKNKGNDKDD 106

136 -YYVYLKPGSKRKNLR---TKQQLAEQVAKGTKEAKK-----GLAQVA 175

10/ KXIVPKPESVEKDLKEMELKEKEFIKQHLKDYEERKEKERNWILRSLRDKLREIEQLE 166

```

1/8 HLSKEEVAANVEAK-----RQGRYTTDDGYIFSPTD-----IIDDL 211

```

18 / KUNAQLSALNELKERRASRRPMMVKMORGKDEVDENIKKYDDEQAENGTKDEEIKDK 226

212 GDAI--LVF---HGNNHYHLPKKDLSPSELAAQAQYWSQKQGRGARPSDYRPTPAPAPGR 266

22 / 0001EELVEINF10MRENALGELD-EYEEERYEKKRYLLKEDGEG-----DLK----- 272

20 KNAE IFDVI ENFGGSHQPFUNGGIHPAPPRPNDA S QNKHQ RDEFGKTI EKELLDQLHRDL 326  
::: : :

2/3 D V E E N E E E I G I G F K E A F F I I N K L E V K K R K N K E Q N A L E K E D K E K K L I A A E E P D D E K K I K L 329

-----EMELADRIEAG 381

SECRETARY OF THE ARMY AND NAVY DEPARTMENT  
WASHINGTON, D.C.  
JAN 10 1968

POINT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
POINT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

T / h	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0																			

[illegible][illegible][illegible][illegible]

THE UNIVERSITY OF CHICAGO

Case	Age	Sex	Occupation	Duration of illness	Site of lesion	Pathological changes	Microscopic findings	Diagnosis
1	25	Male	Teacher	1 year	Brain	Chronic	...	...
2	30	Female	Housewife	6 months	Brain	Chronic	...	...
3	35	Male	Engineer	3 years	Brain	Chronic	...	...
4	40	Female	Teacher	1 year	Brain	Chronic	...	...
5	45	Male	Engineer	2 years	Brain	Chronic	...	...
6	50	Female	Housewife	1 year	Brain	Chronic	...	...
7	55	Male	Teacher	6 months	Brain	Chronic	...	...
8	60	Female	Housewife	1 year	Brain	Chronic	...	...
9	65	Male	Engineer	2 years	Brain	Chronic	...	...
10	70	Female	Housewife	1 year	Brain	Chronic	...	...

633 POLYMERIZATION OF VINYL MONOMERS

[illegible]

677 DVI AKETEVILVDCCTDDEDCCT DETWOCY CO.ETVOCETI

```

703 TRF AK-----CKEVENIRABRENT REVENIINT A-----WCTIIMWIRABRENT REVENIINT- 740
      |||      :|  ||::|||  :::||  |::||  :::||
      .  :  :|  ||::|||  :::||  |::||  :::||

```

735 BREK00ADKSNEN00DSEACKEEEFECN 763

```

750 AKGKEVTNKAHEN-----I EEEVNEED 770
      | | : | : | | | | | | |

```



FILING DATE: 01-SEP-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US94/09943  
 FILING DATE: 01-SEP-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: GATES, EDWARD R.  
 REGISTRATION NUMBER: 31,616  
 REFERENCE/DOCKET NUMBER: L0461/7003  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-720-3500  
 TELEFAX: 617-720-2441  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2465 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-100-804-3

Query Match 3.0%; Score 133; DB 3; Length 2465;  
 Best Local Similarity 19.9%; Pred. No. 0.16;

Matches 176; Conservative 114; Mismatches 322; Indels 272; Gaps 44;

QY 9 GSAVAILAHIGSYOLGRHM---GSATKDNDIAIIDSCKAKAPKTKNTMD-QISAE 64  
 DB 1422 GQVYHLLLEK--GQSPFSKENVHTPOCTLSIDNA--GQGPKEKVKTKTVQVDSFEVE 1476  
 QY 65 EGISABQIVKIKIDGVYSHGDHYFYNGKYVDAIIESELMTPNFRFKOSDINEL 124  
 DB 1477 ENTFEVKLFKNSGSGFSPSRED--NLIPBQINASTVRYKKLFGAGPAESGKIDVDVI 1534  
 QY 125 LDGVIVKVNNTYVYLKPGSKRNINR-----KQO 154  
 DB 1535 L-----KVNQ--ASLKGISQDEVIKALRGAPEVELLCRPPGVLPETIDALLPLQS 1586  
 QY 155 IAEQVAKGTKEAKEKGLAIVAHLSKEEVAAVNEAK-----RGRRTT-----DQYITS 203  
 DB 1587 PAQVLRNNSKSDSQPCVSDSTSDENMSDKSKCKCKSPSRDSTSDSSGSGEDLVTA 1646  
 QY 204 PVIDIDLDGAVY-----VPHGNHYIIPKDLSPSELAAQVWSOKOGRGAPSDY 256  
 DB 1647 PANISSTWSALHOTLSNMVSOAQSHNEAPK--SQEDTICTMFEYPOKIKNPFEDS 1703  
 QY 257 RPPAPAPGRKRRKPIPDVYPRNGQHPONGGYHAPRPPNDASONKHORDEPKGTEKE 316  
 DB 1704 NPSPLP-----PDMA--PGOSTPOS-----ESASSSMQYHIIHISEPTROE 1745  
 QY 317 LLDOLRLDLKYRVHEDDLIFE-PTQVTKSN--AFGYVVRPGDHIIIPRQSLPLEME 373  
 DB 1746 NMPPL-KNDE-NHLEDFELEVEELLITLIKSEKASLGFTVTKGN-----Q 1788  
 QY 374 LADRYLAGOTEDDDSGSDHKSPPSK-----EVHTFELGRIKAYGK-----LDGK 419  
 DB 1789 RIGCYVHADVTPAKSDGRILKPGDRILIKVNDVTMTHTDANVLLRASKIVRLVIGAV 1848  
 QY 420 PVDTSAYVFSKESHSYVSKSGVTAKHGDHFYIGFGELEQYELDEVANWVAKQADEL 479  
 DB 1849 PRITQNTNVAESPATGH-----KLTGNKEEL 1873  
 QY 480 AALDDEQGEKRPLEFTKKVSRKVTKDGKVGYMPKDGKDYFARQDLDLQIAAEQL 539  
 DB 1874 GSSL--CGGHSLSLYOYVYIS-----DLNPRVAIEGNL 1905  
 QY 540 MLKDKHNYRDIYDGT---EPRLAVDSSLPMHAGNATYDTGSSFVYIPIHDIHVVPYS 596  
 DB 1906 QLDLDVHYNGVSTQGMTELEVNRALDM-SLPSVLKATR-----NDLPVPS 1953  
 QY 597 WLTRQDIATIKVVMQPEVRPDIWSPGHEESGYI-PAVTPDLKRAKAPMNOIHSAE 655  
 DB 1954 --KRSVSA-----PKSTGNGSYVSGSQPALTPNDSFSTV-----AGEE 1993  
 QY 656 VOK-ALAEGRFATPDGYIF--DPRDVLAKETPVWKRDSFSIPRADGSSILRTINKSDLQA 712

DB 1994 INEISYPKKCT---YQKSPNLTLPRESYIOEDDI-----DDSOEAEYISLLOV 2045  
 QY 713 ENQAOELLAKRN-AGDA-----TDTD---KPKRQADKSN--- 745  
 DB 2046 D-EEQNLNLNENNAAGDSGPGTLKMGKLSBERTEDDCDPSLPYEFTEATKNGCEE 2104  
 QY 746 --ENQOPEA--SKEEKESD-----FIDSLPDGDLRATED 780  
 DB 2105 YCEKYSLSLQKFOEKKTDDELTWGNELP---IERTHED 2145

# RESULT 9

US-08-296-791-3  
 Sequence 3, Application US/08296791  
 Patent No. 6245337

GENERAL INFORMATION:  
 APPLICANT: St. Gene III, Joseph W.

APPLICANT: Falkow, Stanley  
 TITLE OF INVENTION: Haemophilus Adherence and Penetration

TITLE OF INVENTION: Protein  
 NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:  
 ADDRESS: Flehr, Hohnbach, Test, Albritton & Herbert

STREET: 4 Embarcadero Center, Suite 3400  
 CITY: San Francisco

STATE: California  
 COUNTRY: United States

ZIP: 94111-4187

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/296,791  
 FILING DATE: 25-AUG-1994

CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:

NAME: Trecartin, Richard F.  
 REGISTRATION NUMBER: 31,801

REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989  
 TELEFAX: (415) 398-3249

TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:

LENGTH: 1541 amino acids  
 TYPE: amino acid

TOPOLOGY: unknown  
 US-08-296-791-3

Query Match 3.0%; Score 131.5; DB 4; Length 1541;  
 Best Local Similarity 17.5%; Pred. No. 0.1;

Matches 153; Conservative 111; Mismatches 294; Indels 315; Gaps 37;

QY 44 DSKGAKA-----PKTNKTDQISAEQISAEQIVKTKDGYVTSQDHYHFNK 95  
 DB 462 DNKGSIAKYGDDTVILKQDTNNGSGHAFASVGTISGRSTLVNDKQYDPNSIYFGFRGR 521  
 QY 96 -----VPYDAI-----ISEELMTDPN-----YRFKSDVIN- 122  
 DB 522 LDNGNSLTFDHIINIDGARLVNHNMTNASNITITGESLITDPNTTPYNIDAPDEDNP 581  
 QY 123 -----EILDG---YVYKNGNYVYLKPGSKRNK-----RTKQOIAQVA 160  
 DB 582 YAFRRIKDGQLYLNLLEYVYALRKGASTRSELPRKNGSGENMLYMGKTSDEAKRNV 641  
 QY 161 K-----GTRKAEKGLAIVAHLSKEEVAAVNEAKRQGRYTTDDGYIESPTD 206  
 DB 642 NHINERNNGFNGYFGEERCKNNGNINLVTEFKGSE-----QNFLLTGC-----TN 687



QY 740 -----OAKSNEQOPESEAKESDEDFIDS 767  
 Db 1137 ETVOPOAEAPAREN-DPTVNIKEPOSQNTTADT 1168

## RESULT 11

US-09-514-302-4  
 ; Sequence 4, Application US/09514302  
 ; Patent No. 6338959  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HATADA, Yuji  
 ; APPLICANT: IGARASHI, Kazuaki  
 ; APPLICANT: OZAKI, Katsuya  
 ; APPLICANT: KAWAI, Shuji  
 ; APPLICANT: ITO, Susumu  
 ; TITLE OF INVENTION: GENE FOR ENZYME HAVING BOTH ALKALINE PHOSPHATASE AND  
 ; FILE REFERENCE: 2173-105P  
 ; CURRENT APPLICATION NUMBER: US/09/514,302  
 ; CURRENT FILING DATE: 2000-02-28  
 ; EARLIER APPLICATION NUMBER: 08/952,084  
 ; EARLIER FILING DATE: 1997-11-10  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 4  
 ; LENGTH: 893  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus sp.  
 US-09-514-302-4

Query Match 2.9%; Score 129; DB 4; Length 893;  
 Best Local Similarity 18.7%; Pred. No. 0.071;  
 Matches 131; Conservative 74; Mismatches 218; Indels 276; Gaps 33;

QY 1 VKTGYGYSVAAILLATHIGSYQLGKHHMGSAT-----KDNQIAYIDDSKGAAPK 53  
 Db 90 IDEMYAYDGKLGAE-----HEDGTATLKWSPKADNVSVLYD---KYDQNE 134  
 QY 54 TNKTMDOISAEIGISAEQIVKITT--DQGYVTSBGDHYHFNKGVYDAIISELLMTDP 111  
 Db 135 VVDITENYKGRGWS---VKLTKNVTGLDLSKGYHH----- 169  
 QY 112 NYFRKOSDVINELIDGYIVKNGVNYVYLKPSGKRKNIRTKQIAEQVAKGTKEAKER-- 169  
 Db 170 -YEITHGDTNLTALDPY-----AKSMAAMNNEAGDKVG 201  
 QY 170 --GLAQVAHLKSEEVAAVNEAKRGRYTTDDGYIFS--PTDIIDD--LGDATVLPHGNYH 223  
 Db 202 KAAIVDIGSIGPE---LDYADIPGFEKREDITIIYEVHVDFTSPNIGEDLKAQGTFA 257  
 QY 224 HYIPKDLSPSELAAQA-----YWSQKQGRGAPSDYKRPAPARPKKAPIPDVT 275  
 Db 258 SEVEKLDYI-QELGVTHIQLLPWMSYIFSNEFESGERMLEYASTGT----- 302  
 QY 276 PNPQGHQPDN-----GGYHPPRRPNDASQNKHODEFKGTFKELLQDLHR-----L 324  
 Db 303 -NYNMGIDPHNFTSLSGMSENPEP-----ELRIKEFKNLINETHKRDGMGYVL 350  
 QY 325 DLKRYHVEEDGLIFEPYQVIXSNAGYVVPBGDHYHIIIPRSQLSPELMELADRYLAGOTE 384  
 Db 351 DVVFNHTAQ-----VHIFEDLVP--NYNHEM----- 374  
 QY 385 DDDSGSDHSKPSDEKVTYHTEFLGRKIKAYGKGLDGKPYVTSDAVFSKESI--HSVDKSGVT 443  
 Db 375 -DADGTPT-----VHIFEDLVP--NYNHEM----- 413  
 QY 444 AKHGDFHNYIGGELE--QYELDEV-----ANWYAKAGADDLAALQOE----- 486  
 Db 414 --DGTFRDMGHDABESQIAFDEKALNPNIWIGESWVTFAGDEGPVQADQOMNOY 471  
 QY 487 -----QGEKPELFDTKKYSRKVTKDGKVGKMPKDG---KDYFARDDQLDLTQ 531

Db 472 TEAVGSFSDPEFNEELKSGFSGSGQPRFTT--GGAVWVQOIFPNIKAQPHNFADDPGVQ 530  
 QY 532 IAFADQELMLDKKHRYRDIYDTGIEPRLAVDVSLPMH-----AGNA----- 574  
 Db 531 YTEADNLTLL-----YDVIAOSIKKDPLEAENDLEIHKRIKRGAMAVLTSQGTAFLLA 583  
 QY 575 -----TYDF-----GSSFVTPHIDH 589  
 Db 584 GGEFRTKWRAPATEARYKSTYMTDADGNFVYPIFIH 622

## RESULT 12

US-09-514-302-2  
 ; Sequence 2, Application US/09514302  
 ; Patent No. 6338959  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HATADA, Yuji  
 ; APPLICANT: IGARASHI, Kazuaki  
 ; APPLICANT: OZAKI, Katsuya  
 ; APPLICANT: KAWAI, Shuji  
 ; APPLICANT: ITO, Susumu  
 ; TITLE OF INVENTION: GENE FOR ENZYME HAVING BOTH ALKALINE PHOSPHATASE AND  
 ; FILE REFERENCE: 2173-105P  
 ; CURRENT APPLICATION NUMBER: US/09/514,302  
 ; CURRENT FILING DATE: 2000-02-28  
 ; EARLIER APPLICATION NUMBER: 08/952,084  
 ; EARLIER FILING DATE: 1997-11-10  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 1938  
 ; TYPE: PRT  
 ; ORGANISM: Bacillus sp.  
 US-09-514-302-2

Query Match 2.9%; Score 129; DB 4; Length 1938;  
 Best Local Similarity 18.7%; Pred. No. 0.23;  
 Matches 131; Conservative 74; Mismatches 218; Indels 276; Gaps 33;

QY 1 VKTGYGYSVAAILLATHIGSYQLGKHHMGSAT-----KDNQIAYIDDSKGAAPK 53  
 Db 1135 IDEMYAYDGKLGAE-----HEDGTATLKWSPKADNVSVLYD---KYDQNE 1179  
 QY 54 TNKTMDOISAEIGISAEQIVKITT--DQGYVTSBGDHYHFNKGVYDAIISELLMTDP 111  
 Db 1180 VVDITENYKGRGWS---VKLTKNVTGLDLSKGYHH----- 1214  
 QY 112 NYFRKOSDVINELIDGYIVKNGVNYVYLKPSGKRKNIRTKQIAEQVAKGTKEAKER-- 169  
 Db 1215 -YEITHGDTNLTALDPY-----AKSMAAMNNEAGDKVG 1246  
 QY 170 --GLAQVAHLKSEEVAAVNEAKRGRYTTDDGYIFS--PTDIIDD--LGDATVLPHGNYH 223  
 Db 1247 KAAIVDIGSIGPE---LDYADIPGFEKREDITIIYEVHVDFTSPNIGEDLKAQGTFA 1302  
 QY 224 HYIPKDLSPSELAAQA-----YWSQKQGRGAPSDYKRPAPARPKKAPIPDVT 275  
 Db 1303 SFVEKLDYI-QELGVTHIQLLPWMSYIFSNEFESGERMLEYASTGT----- 1347  
 QY 276 PNPQGHQPDN-----GGYHPPRRPNDASQNKHODEFKGTFKELLQDLHR-----L 324  
 Db 1348 -NYNMGIDPHNFTSLSGMSENPEP-----ELRIKEFKNLINETHKRDGMGYVL 1395  
 QY 325 DLKRYHVEEDGLIFEPYQVIXSNAGYVVPBGDHYHIIIPRSQLSPELMELADRYLAGOTE 384  
 Db 1396 DVVFNHTAQ-----VHIFEDLVP--NYNHEM----- 1419  
 QY 385 DDDSGSDHSKPSDEKVTYHTEFLGRKIKAYGKGLDGKPYVTSDAVFSKESI--HSVDKSGVT 443  
 Db 1420 -DADGTPT-----VHIFEDLVP--NYNHEM----- 1458

OY 444 AKGHDEHYTGFELE--QYELDEV-----ANWYKAKGOADELAAALDOF----- 486  
 Db 1459 --DGFREPMGDHDAESIOIAFDEAKKLNPNIWIGEGWTFEAGDEBPVOAADOOMQY 1516  
 OY 487 -----OGKEKPLFDTKKVSRKVTGDKGVMMPKDG-----KDYFARDOLDLTQ 531  
 Db 1517 TEAVGSFDEFRNELKSGFSGEGPRFIT--GAVNVQOIFUNIKAFPHNFADDPGQVYQ 1575  
 OY 532 IAFABOELMLDKKHRYRYDVTGIEPLAVDVSSLPNH-----AGNA----- 574  
 Db 1576 YIEAHNDLTL-----YDVIAQSIKDPREIAENDLEIHKRIRYGNMAMVLTSOGTAEFLHA 1628  
 OY 575 -----TYDF-----GSSFYIPIHDH 589  
 Db 1629 GQEGFRTKQWRAPATEARYKSTYMTDADGNPFYIPIFH 1667

## RESULT 13

US-08-867-941-19  
 ; Sequence 19 Application US/08867941  
 ; Patent No. 5977337  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Loosmore, Sheena M  
 ; APPLICANT: Du, Run-Pan  
 ; APPLICANT: Wang, Quijun  
 ; APPLICANT: Wang, Yan-Ping  
 ; APPLICANT: Klein, Michel H  
 ; TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA  
 ; NUMBER OF SEQUENCES: 67  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Slim & McBurney  
 ; STREET: 6th Floor, 330 University Avenue  
 ; CITY: Toronto  
 ; STATE: Ontario  
 ; COUNTRY: Canada  
 ; ZIP: M5G 1R7  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/867,941  
 ; FILING DATE: 03-JUN-1997  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Stewart, Michael I  
 ; REGISTRATION NUMBER: 24,973  
 ; REFERENCE/DOCKET NUMBER: 1038-681 MIS:j)b  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (416) 595-1155  
 ; TELEFAX: (416) 595-1163  
 ; INFORMATION FOR SEQ ID NO: 19:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1076 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; US-08-867-941-19

Query Match 2.9%; Score 128.5; DB 2; Length 1076;  
 Best Local Similarity 20.1%; Pred. No. 0.1;  
 Matches 141; Conservative 87; Mismatches 244; Indels 231; Gaps 40;

OY 3 KTYGYGSAVAAILLATHISYOLGK-----HHMSATKNOIAYI-----DD--SKGK----- 48  
 Db 148 KNYAAGGAINLEY-ENVSVEISKANSSEYSGALSSVAFVYKTFADDIIRKDGKMGV 206  
 OY 49 ----AKAPTKNKTMDOISAEEGISAEQIVYKTIID--OGYVTSCHDHYHFYNGKPYDAI 101  
 Db 207 QTTATATASKNMAMVNSVAAGAKGSGISGLIITYDRGQYER-KAHD--AYGSGSDFDA 262

OY 102 ISEELMTDPNVEFKOSDVINEILLDGYVIKNGNYVYLKPK-----SKRNKPIRTKQIA 156  
 Db 263 VA-----TIDPNNR--TELLANEC-----ANGNEEACAGAGQTKLQAKPTVRBKVNVK 309  
 OY 157 EOYAKG-----TKAKEKGLAOUVAHLS--KEEVAANVNAKQGRYTTDDGIIESPTDII 208  
 Db 310 DYNGPRLIIPNPLTODSKSILLRPGYOLNDKHYGGVYEITFO--NVAOMDKTV----- 361  
 OY 209 DDGDAVILVPHGNHYHYIRKKDLSPSEL--AAAOAYWQOKGARGARPSDPTPAPARG 265  
 Db 362 ----PAYLAH-----DIEKRLSNHAQANQY--QGNLGERIAD----- 396  
 OY 266 RRRAPIDYTPNPGOHDPDNGGYNHAPPRPDASQKKHQREE-----KGTFK 315  
 Db 397 -----TIGPDSGYGINAHGVFY-----DEKHQKDLGLELYYDSKGENKWD 439  
 OY 316 ELIDDLHRLD-----KYRHVERDGL--IEFTQYIKSNAPGYVPH----- 355  
 Db 440 DVHVSYDKODITLRSQLTNHCSTYPHIDKCTPYVKNKPSYKEVDNNAYKQHNLIKAV 499  
 OY 356 -----GDYHILIPR-----SOLSPLEMLADRYLAGOTEDDSSGSHSKPSDKEV 400  
 Db 500 FNKKMALGSTHHHINIQGYDKFNSSLRVEYRFLA-----THQSYOKLDYTPSPNP-- 550  
 OY 401 THFFLGRIRKAYGKLDGKPYDTSDAIVYSKESIHVSVDKSGYAKKGDHFIHIGFELEQ 460  
 Db 551 ----LPDKFKPI--LGSNNKPI--CLDAYGYCHD-----HPQACNAKSNYONFAIKKIGIO 599  
 OY 461 YELDEVANWYKAKGOADEL--AAALDO--EOGKEKPLFDTKKVSR-----KTKDKGVGY 511  
 Db 600 Y-----NOKNTNDKIDYQAIIDQYKONNSITLKPREKIKOSLGEQKYKIDELGF 650  
 OY 512 MPRKD-----GNDYFARDOLDLTQIAFAEOELMLDKKHRYRYDIV 552  
 Db 651 KAYKDLNEMAGMTNONSQONANKGTDNIIYQ-----PNQATVYKDK--CKYS-- 696  
 OY 553 DTGIEPLAVDVSSLPNHAGNATYDGGSFVITPHIDHIVPY 595  
 Db 697 ----ETNSVADCSSTPRH1-----SGDNFIALKDMTINKY 729

## RESULT 14

US-09-074-658-19  
 ; Sequence 19 Application US/09074658  
 ; Patent No. 6184371  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Loosmore, Sheena M  
 ; APPLICANT: Run-Pan Du  
 ; APPLICANT: Quijun Wang  
 ; APPLICANT: Wang, Yan-Ping  
 ; APPLICANT: Klein, Michel H  
 ; TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA  
 ; NUMBER OF SEQUENCES: 78  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Slim & McBurney  
 ; STREET: 6th Floor, 330 University Avenue  
 ; CITY: Toronto  
 ; STATE: Ontario  
 ; COUNTRY: Canada  
 ; ZIP: M5G 1R7  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/074,658  
 ; FILING DATE: 08-MAY-1998  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Stewart, Michael I  
 ; REGISTRATION NUMBER: 24,973  
 ; REFERENCE/DOCKET NUMBER: 1038-795

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1163  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1076 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-09-074-658-19

Query Match 2.9%; Score 128.5; DB 4; Length 1076;  
 Best Local Similarity 20.1%; Pred. No. 0.1;  
 Matches 141; Conservative 87; Mismatches 244; Indels 231; Gaps 40;

3 KTYGIGSVAAILLATHISGYOLG---HHGSAKDKNOIAYI---DD-SKCK-----48  
 148 KNAAGGALINELEY-ENVSVELSKANSSEYSGSLSSVAEYKTTADDTIKGDKMGV 206  
 49 ---AKAPTNTKTMQDISAEIGISAEQIVKTTD---QGYVTHSGDHHFHYNGKVPYDAI 101  
 207 QTKTAYASNNMAVNSVAAGKAGSFGSLITDRRGQY-KAHD-AYGSGSFDFA 262  
 102 ISEELMTPDNTREFKOSDVINELDSYIKVGNYYIKPG-----SKRKINIKQOIA 156  
 263 VA-----TTDPNNR--TFLLANEC-----ANGNEACAAGQTKLQAKPTVWRDKVNVK 309  
 157 EQVAKG-----TKRAKEKGLAOVAHLS-KEEVAVANEAKRGRTTDDGYTFSPDTII 208  
 310 DYTGRRLIPNPLTQSKSLILRPGYQNDKHYGVGYEITKO-NVAMQDKV-----361  
 209 DDGDAVYLVPHGNHYIYIKKDLSPSEL---AAQAYWSQKQGRGAPSDYRTPAPAPG 265  
 362 ---PAYIAVH-----DIEKSRLSNHAQANGYV-QGNNLGERLRD-----396  
 266 RRAKAPDYTPNPGCHQDPNGCYHAPRPAPRDASQNKHQRDEF-----KGKTFK 315  
 397 -----TIGPDSGYGINAHGVY-----DEKHQKRLGLEEYVYDSKGENKMPD 439  
 316 ELIDDLHLRLD-----KYRHEEDGL-IFEPYQYKSNAGYVVPH-----355  
 440 DRYVSYDKODITLRSOLNTHTCSYRPHIDKCTPRVKNKFSYKEVDNNAKQNHLLIKAV 499  
 356 -----GDHYHIIIR-----SOLSPLEMLADRYLAGOTEDDSDGSHKSDKEV 400  
 500 FNKKAALGSTHHILNQGYDKFNSSLSRYEYRLA-----THOSYQKLDYTPPSNP- 550  
 401 THTFLGRIRKAYGKGDGKPYDTSDAVYFSKESHSYVDSKGVTAKHGHFHYIGSGLEQ 460  
 551 -----LPDKFKPI-LGSNNKPI-CLDAYGCGHD-----HPQACNAKNSYIYNFAIKKGI EQ 599  
 461 YELDEVANMVKAKGQADEL--AAALDQ--EQQEKRPLEDTKKVSR-----KYTKDGKYG 511  
 600 Y-----NOKTMTDKIDYQAIIDQYDKQNPSTLKFPEKIKGSLQEKYKIDELG 650  
 512 NMPKD-----GKDYFYARDQDLQIQAEBELMLKHKHRYIIV 552  
 651 KAYKDLREMGAGTWDNSQQAANKSTDIYQ-----PNOATVYKDK-CKYS- 696  
 553 DTGIEPRLAADVSLPMHAGNATYDTGSSFVPIPHIDHIVVY 595  
 697 ---ETNSYADCSSTPRHI-----SGDNYFTALDKDMNTINKY 729

RESULT 15  
 US-08-663-112-2  
 ; Sequence 2, Application US/08663112  
 ; Patent No. 5849503  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MAGATSUMA, Masako  
 ; APPLICANT: KURITA, No. 58495031ko  
 ; TITLE OF INVENTION: MUTANT PROTEINS OF HUMAN DNA  
 ; TITLE OF INVENTION: TOPOISOMERASE I

NUMBER OF SEQUENCES: 7.  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
 ADDRESSEE: Dunner L.L.P.  
 STREET: 1300 I Street, N.W.  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20005-3315  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/663,112  
 FILING DATE: 26-NOV-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Einaudi, Carolyn P.  
 REGISTRATION NUMBER: 32,220  
 REFERENCE/DOCKET NUMBER: 06609.1488-00000  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-408-4400  
 TELEFAX: 202-408-4000  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 765 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-663-112-2

Query Match 2.9%; Score 127; DB 2; Length 765;  
 Best Local Similarity 20.1%; Pred. No. 0.082;  
 Matches 137; Conservative 90; Mismatches 240; Indels 214; Gaps 34;

141 KPGSKRNKIRTKQOIAEQVAKTEKEKGLAOVAHLSK-----EEVAAVNEAK 169  
 42 KDREKSKNSNEHDSKKHKEKTKHKKOSSKHKDKHKDRKREKREKRAASDAK 101  
 190 RQGRYTTDDGYTFSPDTIID-IGDAVLPVPHGNHYIYIKKDLSPSELAQAQVWSQKQ 248  
 102 I--KKEKENF--SSPQIKEDPEDDGYFVP-----PKEDIK-----LKRP 139  
 249 RGAPRSDYRTPAPAPRRKAPRPDY--TPNPGCHQDPNGCYHAPRPAPRDASQNKHQRD 307  
 140 RDEDDVDIKPKKIKETEDTKKEKKRKLLEEDGKLKPKRKDKKQVPEPNKKRKKKEE 199  
 308 EFKGTFKELDLHLRLDLYRYHVEEDGLIFEPY-----QYIK-----345  
 200 EQKWKWME--ERYPRGKWKPLEHKGYPAPRYEPLPENVKFYTGKWKMLSPKAEV 256  
 346 SNAIGYVVPH-----GDHYHIIIPRSQSPLEMLADRYLAGOTE 384  
 257 ATFFAKMLDHEYYTTELKERNFKQMRKREMTNEKNII--TNLSKCDPTQMSQYFAQTE 314  
 385 DDDGSGSHSKRSDKEVYHTFLGRIRKAYGKGLDGKPYDTSDAVYFSKESI--HSVDKSGV 442  
 315 ARKMSAEELIKIKEENEKLL-----KEYG-----FCIMDNH--KERIANFKIEPPGL 360  
 443 TAKHGDHFIYIGFELBOYELEDE--VANWVKAKGQADELAALDQEGKPKPLFTKRYV- 499  
 361 FRGRGNH--PKMGLMKRLIMPEDIIINCSK-----DAKVPSPPGHKKKKEV 405  
 500 -SRKVT-----KDKKGYVM-----FKDGKDYFYARDQDLQIQAEBELML 541  
 406 HDNKVTVLVSTENIQSITYIMNPSSRIKGERDKWQKYEYAR-----RLKKCYD 455  
 542 KDKKHRYDYIVDNGIEPR-----LAVDVSSLPNHAGN-----ATYDGS--SFVYPIHD- 588  
 456 KIRQYEDMKSKEMKQRAVAVALYFIDKLALAGNKEGEGETADTYGCCSLVYEHINLH 515







QY	469	IDIQNNLKIOTISFLSKKRLTLVAADNGIEDIRPGQAPNKFIVLSNNKISDLSPIAS	528
QY	158	LDISSNNKVSDDSYALKTLNIESCLATNNOISDTPELITNLBELSLNGQDGLTAS	217
QY	529	LHQQLIHDINNOITDLSPVSHKESLTIVYDLSRNADVDLTLQA-PKLETLAMVDKVS	587
Db	218	LTNLTDLDLANNOISNLPAISGLKRTLELKGANOISINISPLGLTALTNLEINOLED	277
QY	588	LDPLKNPNLSTLSITNAQOLESIGTFASSIVYVEAEQNOIKSLYVKDQGSITFLPDVT	647
Db	278	ISPLSNLKNLTLYLLYLRNNISDISIPVSSLRKQPLFFYNNKVSDDVSLANLTJNNMLSAG	337
QY	648	GNOLTSLEGVNNFFALDILYSKKQOLTN-----VNSKRP--KTYTNIDISHNNIS---	656
Db	338	HNOISDLPPLANLNRITOLGLINDQAMTNAPYNNKANSIPTKYNNVIGALIAPATYISDGC	397
QY	697	---LADLKLNEQHIPEALAKNFP---AVYEGSNGCSTAEB--KAAMATK---AKESAQ	744
Db	398	SYAEPDITWNLPSYTNESYTFESQPVITGKOTTFESGTVQPLKALFNAAKFHVQDEKETT	457
QY	745	EASESHDYNNHNTYEDEBEGH	764
Db	458	EV-EAGNLTLEPAKPVKEGH	476

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

Db      213  SOLSYSTASJONN-----TQSVAKGSTS---KPANSEMJOSLTEL--YDSPOQRYS 261
Qy      71  -DDEILLFKOSKILSKTDGIVADHDSHFIFYADLKG--SPFEYLIP--KGASL--- 121
Db      262  ESDGLVF-DPAKIISRIPNGVAIRHGHYIFIPYKSLALEEKIARVPLSGTGVYSTN 320
Qy      122  AKP-----AVAQRAASOCTS-VADPHHHKEFNADIVAEADALGYTRHDDHHYIL 172
Db      321  AKPHEVYSSLSLSSPNPSSLTTSKELSSASDGYFNFKRDIVEETATYIVRHDDHHYIL 380
Qy      173  KSSLISGQAOAKAVATRLPOTSSLSSTATANGIPLGHLPTSGFOFNGOIGVYTKDSI 232
Db      381  KSNQGGQDTLPNNSLAT--PSPSLPIPGTSH-----EKHEEDGCGFDANRIIAEDSGF 433
Qy      233  LVYDHDGHLH 241
Db      434  VMSHGDHNH 442

```

```

RESULT 3
US-08-961-083-56
; Sequence 56, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO.: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 796 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-961-083-56

Query Match          5.5%; Score 223; DB 4; Length 796;
Best local Similarity 19.5%; Pred. No. 7e-09;
Matches 168; Conservative 116; Mismatches 279; Indels 300; Gaps 42

QY      30 TYPIKTKSGRKGTGSKIKIPKSKKTNTNGVGAVDEPTDGGFLITLDSKITLDG 89
       ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db       1 SYELGLVQATWAKENRNVSYDGKQATQKT--EMLTPDEVSKREG--INAEQIYIKITLDG 57
       |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      90 IYVDHGHSHFIYADLKSGPFEEYLPRKGASLAPPAVAORAASOGTSKVADRRHHNYEFNP 149
       |||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

Db 58 YVTHSGDHVY-----YNGKVPYDAIISELL-----MKDP--NKKLD 94  
QY 150 ADIYAEDALGYTVRHDHDFHYLLKSS-----LSGOTQAQAKOYATRLPQTSYLV 198  
Db 95 EDIVNEVKGYYVTKGQYVYLKDAHADNVRTKEINQKQEHSGHREGGPRNDGAV 154  
QY 199 STATANGICPLHPTSGFQFNGQGYVTKDSILVDHGHLPISFADL-----248  
Db 155 ALARSQG---RYTDDGYIFNADIEDGDYIVPHGHYHYIPKNEISASELAABA 210  
QY 249 -----ROG-----GMAYADQYDPAKKAKEAETHQTPELSEKEKEQ 286  
Db 211 FLSGRNLMSKRYTRKNSDNTSRITNM--VPYSNPGTTNTNTSNNSTNSQSGNDID 268  
QY 287 EKLAYL-----AEKLGIDPSTIKRVETQDGLGLEYPHHDAHVL--MLSDIE 332  
Db 269 SLKQLYKLPLSGRHYESDGLVDFDPAQI---TSRTARGAVPHGHYHPIPSQSELE 324  
QY 333 --IGKIDP---DPHAIEHARELEKHKVGMOTLRALGFDEEVLIDIVRTHDAFPPTSNE 386  
Db 325 ERIARIIPLRYRSHNVWPPDSRP-----EOPSPQPTPE 356  
QY 387 KDNMMKEMLATYVIXLIDGSRKDPLOKGLSILPNI-ETLIGTFPIKDISPYLOKRIK 445  
Db 357 PSFGPOP--APMLKIDSNS-----SLVSQLVKRYGEGY-----387  
QY 446 QLMTRTGYVDFEFLDNMPLEGIDISQNNLKDI-SFLSKYKNI--TLVAADNGIEDIR 502  
Db 388 --VEEEKISRYVFAKDLR-----SEYVKLESKLSQESVSHTLVAKKN-----VA 433  
QY 503 PLGO-----LPLN-----KFLVLSNNKISDLSPLASLHLOELHIDNNOITDLSVSHKE 552  
Db 434 PROEFYDKAYNLTEAHKALFYKNGKNSDFQALDKL--LERLNDES-----TNKE 482  
QY 553 SLTVNDSRADVDLTIQAPKLETILVNDTKYSHLDFLKNPNLSLSINRAQSLSG 612  
Db 483 KL--VD-----DLAFLAP-----ITHPERL-GKPN-----505  
QY 613 IEASVIVVEABGNOIKSLYLKDKOSLTFDVTGNQTLSEGVNNFTALDILSVSKNO 672  
Db 506 -----SQIEYTEDEIRIQLADK-----YTTSDGY-IFDEHDIISDEGIA 544  
QY 673 LTVNLSKPKVITYNIDISH--NNISLADIKLEOHIPRAIAKN--PRAVYEGSWANG 727  
Db 545 YVTPHMGH-----SHWICKDSLSD--KEKVAQAQYKEKGLPSPADAVKANP 591  
QY 728 TAKEKAMATKAKESQAESHSYDNNHNTYEDEBGA-----HHHRDK--DDHDH- 776  
Db 592 TGSAAAIYKRVGKRIPLVRLPYVEHTVEYKNGNLIIPKDHVHNIRFAMFDDHTYK 651  
QY 777 -----EHEDE 781  
Db 652 APNGYTLDELATIKYVEHPDE 674

RESULT 4  
US-08-961-083-66  
; Sequence 66, Application US/08961083  
; Patent No. 6159469  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
; COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/961,083  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 763 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-961-083-66

Query Match 5.0%; Score 204; DB 4; Length 763;  
Best Local Similarity 19.8%; Pred. No. 1.9e-07;  
Matches 171; Conservative 125; Mismatches 314; Indels 254; Gaps 43;

QY 41 GMTSMKIKPIKSKKTNTKNGVAGVDFPTDDGFLITKDSKILSTKDGIVVDHGHSHF 100  
Db 29 GOKAENLTPDEYSKREG-----INAEQYVIKITDGYVTSGHGDHY 70  
QY 101 IFYADKSGPEFLIPKGSALAPVAGRAASQGSTKADPHNHEFPADIVADALGY 160  
Db 71 -----YNGKVPYDAIISELL-----MKDP--NKKLDSDIVEIKGY 107  
QY 161 TVRHDHDFHYLLKSSLSGOTQAQAKOYATRLPOTS-----SLVSTATNGIPGLHFP 212  
Db 108 VAKVNGKYVYLKDAHADNIRTKETIKROKQERHNHNSRADNVAARAQSG---RYT 163  
QY 213 TSDGFQFNGQGYVTKDSILVDHGHLPISFADLRGGMAHVAADQYPAKKAKEPAET 272  
Db 164 TDDGYIFNADIEDTGDYIVPHGDHNYIPKNEISASELA-AAEAVMNKGQSRPSS 222  
QY 273 HQ-----TPELSER-----EKEYOEKLYAEKLGIDPSTIKRVETQDGL 313  
Db 223 SSYMANPAPQRLSENNHNLTVPTYHONOGENTISLRELVAKPUSERVEES-DGLIEDPA 281  
QY 314 -----GLEYPHHDHVAHL--MLSDIE--IGKIDP---DPHAIEHARELEKHKVGM 356  
Db 282 QLTSTARGAVAPHGNHNFIPYEQMSELEKRIARITPLRYRSHNVWVDSRPPQ-----336  
QY 357 DTLRALGFDEEVLIDIVRTHDAFPPTSNEKODPNMKEMELATVIXLIDGSRKDPLOKGL 416  
Db 337 -----SPQSTPEPSPPOAPNPQP--APSNIIDKLVKEAVRKVD 376  
QY 417 SLPLNLETLGIG-FPIKDISPV-----LOPKIKLO-LMTGTGV--TP-----456  
Db 377 GYV--FEENGVSRTYPAKDLAETAAGIDSKLAKGESISHKLGAKKIDLPSSDREYVKA 434  
QY 457 YFELDNMPLEGIDISQNNLKDISFLSKYKNIITLVAADNGIEDIRPLGOLPNIKFLVLS 516  
Db 435 YDLARHQ---DLDDMKGRQVDF-----EALDNLLEKRD-----VX 469  
QY 517 NNTI-----SDLSPLASLHLOELHIDNNOITDLSVSHKESITLVVDLSRNADVLATLQA 572  
Db 470 SDKVLVXDILAFAPIRHPRERLGRPNAQIITYTD-----DEIOVAKLAGKYYTDEGYTFD 524  
QY 573 PKLETILVNDTKV---SHLDPLKNNPNLSISINRAQSLSE-GIEASSV---IYRVE 623  
Db 525 PDLITSDGDAVYVTHMTHSHWIKKO-SLSEAEARAQAQVYKKEGLTPPSTDHODSGNTE 583





```

QY 490 LYAADNGIEDIRP--LCQLPRLKFLVLSNNKIS-----DLS----- 524
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 219 ELDLSDNALNGSIPASLGNMNLFLFLYGNOLSGSIPETIGYLSRLTYLDLSENALNGS 278
QY 525 ---PLASHOLOELHIDNNQITDLSPP--VSHKESLTIVDLSRNA-----DVDLATLQAPK 574
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 279 IPASIGNLNLFLFLYGNOLSGSIPETIGYLSRLTYLDLSENALNGSIPASLGNKMLLS 338
QY 575 LETLVNDTKVSHDLFLKNNPNLSL-----SINRAOLOSLEGEIASSV 618
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 339 RLNLVNNOLSGSIPASLGNLNLMLLYNNOLSGSIPASIGNLNLMLLYNNOLSGS 398
QY 619 IYRVAEAGNOISLYLAKKO-----GSLTFLDYNGNOLSL--EGVNNFTALD 664
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 399 IPASIGNLNLRLTYLNNOLSGSIPETIGYLSRLTYLDLSSNNGTIPASFGMSMLA 458
QY 665 ILVSKNOITVNLKSPN-----KVTNIDISH-----NNISLADIKLE- 704
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 459 FLFLYENOLAS---SVPEETIGYLSRLTYLDLSENALNGSIPASFGNLNLSRLTYLNNOL 515
QY 705 -OHIPPAI 711
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 516 SGSIPEEI 523

```

## RESULT 9

```

US-09-228-986-72
; Sequence 72, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Neuenhuzen, Niels
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 910
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-72

```

Query Match 4.0%; Score 165; DB 4; Length 910;  
 Best Local Similarity 23.9%; Pred. No. 0.00026;  
 Matches 90; Conservative 63; Mismatches 114; Indels 110; Gaps 17;

```

QY 416 LSLPMLLETIGIFPFIKDISP-----VLQFKLKQLIMKTGVTDYRFLDNNMPQLEG 468
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 171 LCLLPMLICSLGGINNLGTIPDCLGNISLQYLSQGNLOGSVSE---LGRLSQLIV 227
QY 469 IDISQNNLKDI--SLSKYKNLTLYAAADNGIEDIRPQLQDPN-----LKLFLVLSN 517
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 228 LDFGNHLTGCIPISSLSNCTNELLDIGNOL---VGHIPSHLCTKTKTQTLMLYRLGA 282
QY 518 NKISDLSPP--LASLHLOLELHIDNNQITDLSPP--VSHKESLTIVDLSRNAVDLATL 570
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 283 NOLSGSVPSLFFNCTKLOELALPYNOLSGIVPMELCKLTHLQRLP---FGNITFISGNTM 339
QY 571 QAKRLLETIMVNDTKVSHDLFLKNN-----PN 596
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 340 KCPIL--TALSNCSDLOYDISENNFTGOLPFIHGHSKLLYHLDGSLNLAGIIPPAIGN 398
QY 597 LSSLST-----NRAOLOSLEGI--EASSVYRVAEAGNOISLYLAKDKOGS 640
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 399 LSSLTFLNLGRNYFTGISPSLSLIMLQRLBYLNDNNLOGNIPMETGQSLGL----- 452
QY 641 LTLFDVNGNOLSL--EGVNNFTALDILYSKKNOLP---NVNLSKPNKTVTNIDISHNNI 695
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 453 ---LYISGNLNLGKIPDPYANLQOLRYLYLNHOLSGDINANLGKCVNLLL--LDLSYNNL 508

```

```

QY 696 SLADLKLNEQHIEPAIA 712
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 509 S-----GHIEPELA 517

```

## RESULT 10

```

US-09-180-439-3
; Sequence 3, Application US/09180439
; Patent No. 6225532
; GENERAL INFORMATION:
; APPLICANT: Dixon, Mark S
; APPLICANT: Halzanthis, Kostas
; APPLICANT: Jones, David A
; APPLICANT: Jones, Jonathan DG
; TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof
; FILE REFERENCE: 620 - 53
; CURRENT APPLICATION NUMBER: US/09/180,439
; CURRENT FILING DATE: 1998-12-06
; EARLIER APPLICATION NUMBER: PCT/GB97/01249
; EARLIER FILING DATE: 1997-05-08
; EARLIER APPLICATION NUMBER: GB 9609681.3
; EARLIER APPLICATION NUMBER: GB 9619924.5
; EARLIER FILING DATE: 1996-09-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-09-180-439-3

```

Query Match 3.9%; Score 160; DB 4; Length 968;  
 Best Local Similarity 21.3%; Pred. No. 0.00069;  
 Matches 119; Conservative 86; Mismatches 204; Indels 148; Gaps 26;

```

QY 287 EKALYLAE--KLGDIPSTIKRVETQDCKLGEYPPHDAHVLMLSDETIGKIDPPHPAIE 344
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 209 EETIGYLSRLTYLDLSENALNGSIPASIGNLNLMLLYNNOLSGSIPETIGYLSRLTYLDLSENA 260
QY 345 HARELEKIKVMDTL-----RALG-FDEEYIYLDIVRTHDAPTPPPSXPNNMKKE--WL 396
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 261 YKLSITKLSLIGINFLSGSIPASIGNLNLMLSLDLY-----NNKLSGSIPETIGYL 310
QY 397 ATVIKLDLG-----SRKDPLORK-----GLSLPMLLETIGIFFP 431
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 311 RSLTYLDLSENALNGSIPSSIGNLNLMLSLDLYNNKLSGSIPETIGYLSRLTYLDLSENA 370
QY 432 IKDISPVLOFKKQLQMLT-----KTGVTDYRFLDNNMPQLEGIDISON 474
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 371 LNSGIPA--SLGNLNLMLLYNNOLSGSIPETIGYLSLTEL--YLGNNSLNGSIPASIG 428
QY 475 NLKDISLSKXKN-----LTLVAAADNGIEDIRP--IGOLPMLKFLVLSN 517
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 429 NLNNLPMLYLNNOLSGSIPETIGYLSLTELFLGNNSLNGSIPASIGNLNLMLSLTYLN 488
QY 518 NKISDLSPP--LASLHLOLELHIDNNQITDLSPP--VSHKESLTIVDLSRN----- 562
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 489 NOLSGSIPASFGNKNRNLQTLFLSDNDLIGEIFSVCNLSLEVLYMGRNMLKGVPOCLG 548
QY 563 ---ADVDLATLQA---KLETLVNDTKVSHDLFLKNN-----PNLSLSINRAQL 607
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 549 NISDLHLSMSSNSFRGLPSSISNLTLKILDDGGRNNLEGATPQFFGNISLQVFDMMON 608
QY 608 OSLEG-----IEASSVYRVAEAGNOISLYLAKDKOGSITFLDYNGNOL--TSLEGVN 658
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 609 NKLSTGLPTNSIGCSLSLMLHGNELADETPRSLDCKKRLQVADIDQNDLNDFFPWWLG 668
QY 659 NFTAIDLVSKNOL-----TNVNLSPKNKTVTNIDISHNNISLADLKLNEQHIEPAIA 712
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 669 TLPELRVLRITSNKLHCPRISSGAEIMPPDURI--IDLSRNAFS-----QDLPTSLF 718
QY 713 KNFPAVYEGSNVGNCTALE 731

```



```

1  APPLICANT:  LONGAKER, MICHAEL T.
2  APPLICANT:  WHITBY, DAVID J.
3  APPLICANT:  HARPER, JOHN R.
4  APPLICANT:  PIERSCBACHER, MICHAEL D.
5  APPLICANT:  BORDER, WAYNE A.
6  TITLE OF INVENTION:  INHIBITORS OF CELL REGULATORY FACTORS
7  TITLE OF INVENTION:  AND METHODS FOR PREVENTING OR REDUCING SCARRING
8  NUMBER OF SEQUENCES:  8
9  CORRESPONDENCE ADDRESS:
10 ADDRESS:  CAMPBELL AND FLORES
11 STREET:  4370 LA JOLLA VILLAGE DRIVE, SUITE 700
12 CITY:  SAN DIEGO
13 STATE:  CALIFORNIA
14 COUNTRY:  UNITED STATES
15 ZIP:  92122
16
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE:  Floppy disk
19 COMPUTER:  IBM PC compatible
20 OPERATING SYSTEM:  PC-DOS/MS-DOS
21 SOFTWARE:  PatentIn Release #1.0, Version #1.25
22
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER:  US/08/303,238
25
26 FILING DATE:
27 CLASSIFICATION:  514
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER:  US 07/978,931
30 FILING DATE:  17-NOV-1992
31
32 ATTORNEY/AGENT INFORMATION:
33 NAME:  KONSKI, ANTOINETTE F.
34 REGISTRATION NUMBER:  34,202
35 REFERENCE/DOCKET NUMBER:  P-LA 9453
36 TELECOMMUNICATION INFORMATION:
37 TELEPHONE:  619-535-9001
38 TELEFAX:  619-535-8949
39
40 INFORMATION FOR SEQ ID NO:  4:
41 SEQUENCE CHARACTERISTICS:
42 LENGTH:  359 amino acids
43 TYPE:  amino acid
44 TOPOLOGY:  linear
45
46 MOLECULE TYPE:  protein
47
48 US-08-303-238-4

```



RESULT 15  
US-08-190

```

US-08-190-802A-50
: Sequence 50, Application US/08190802A
: Patent No. 5519003
: GENERAL INFORMATION:
: APPLICANT: Mochly-Rosen, Daria
: APPLICANT: Kon, Dotit
: TITLE OF INVENTION: WO-40 - Derived Peptides and Uses
: TITLE OF INVENTION: Theoeof
: NUMBER OF SEQUENCES: 265
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Dellinger & Associates
: STREET: P. O. Box 60850
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94306-0850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/190,802A
: FILING DATE: 01-FEB-1994
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Fabian, Gary R.
: REGISTRATION NUMBER: 33,875
: REFERENCE/DOCKET NUMBER: 8600-0139
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 324-0880
: TELEFAX: (415) 324-0960
: INFORMATION FOR SEQ ID NO: 50:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 603 amino acids
: TYPE: amino acid
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: HYPOTHEITICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: INDIVIDUAL ISOLATE: Insulin-like growth factor bind
: INDIVIDUAL ISOLATE: pro. complex-ital, Fig. 33
US-08-190-802A-50

```

Query Match 3.88; Score 153.5; DB 1; Length 603;  
Best local similarity: 3.28

Matches 104; Conservative 77; Mismatches 174; Indels 91; Gaps 20;

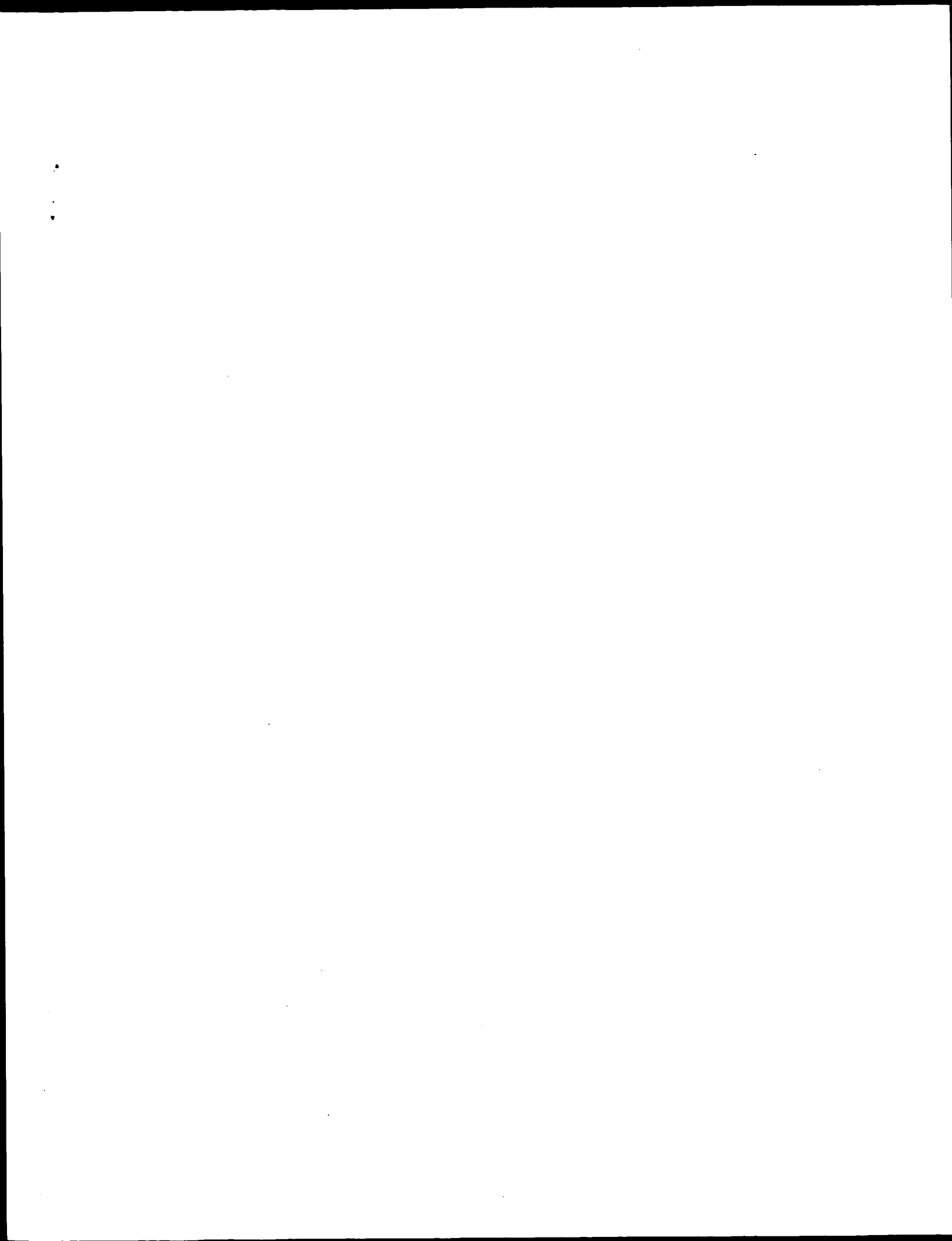
[illegible]

```

Dh      388  LHSLEHSEGLGHVRLHMFAGHSGLRRLFLRDN-----SISIEEQSLAGL-----433
QY      622  VEAEGNQISLVKKNQSSLEFLFDYTGQNLST-----EGVNNFTALDITLSYKNQLTNV 676
Dh      434  -----SELLELDITNRLRHLRPLFOGLGH---LEVLISYQNLTL 473
QY      677  N-SLSPKNTVTINIDISHNNI-SLAD 699
Dh      474  SAELVGLPQRAFWDISINNHETLAE 499

```

Search completed: May 19, 2003, 08:24:47  
Job time : 19.3699 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 07:51:03 ; Search time 227.541 Seconds  
(without alignments)  
335.715 Million cell updates/sec

Title: US-09-645-835A-4

Perfect score: 4080  
Sequence: 1 MKTKKVIIVGLLSSQLTL.....DHDHEDENAKDEQNHAD 792

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 362588 seqs, 96450795 residues

Total number of hits satisfying chosen parameters: 362588

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCCTS\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	292.5	7.2	1039	US-09-884-465A-6	Sequence 6, App1
2	292.5	7.2	1152	US-09-884-465A-379	Sequence 379, App
3	292.5	7.2	1238	US-09-884-465A-381	Sequence 381, App
4	292.5	7.2	1365	US-09-884-465A-382	Sequence 382, App
5	292.5	7.2	1378	US-09-884-465A-378	Sequence 378, App
6	241	5.9	715	US-09-253-088-16	Sequence 16, App1
7	241	5.9	793	US-09-253-088-15	Sequence 15, App1
8	233	5.7	447	US-09-765-272-182	Sequence 182, App
9	224.5	5.5	840	US-09-884-465A-10	Sequence 10, App1
10	224	5.5	838	US-09-884-465A-8	Sequence 8, App1
11	223	5.5	796	US-09-765-272-56	Sequence 56, App1
12	204	5.0	763	US-09-765-272-66	Sequence 66, App1
13	201.5	4.9	2026	US-09-801-368-86	Sequence 86, App1
14	195	4.8	1119	US-09-905-291A-294	Sequence 294, App
15	195	4.8	1119	US-09-902-853-294	Sequence 294, App
16	195	4.8	1119	US-09-907-824-294	Sequence 294, App
17	195	4.8	1119	US-09-907-841-294	Sequence 294, App
18	195	4.8	1119	US-09-904-011-294	Sequence 294, App
19	195	4.8	1119	US-10-028-072-352	Sequence 352, App

20	195	4.8	1119	US-09-906-742-294	Sequence 294, App
21	195	4.8	1119	US-10-121-049-352	Sequence 352, App
22	195	4.8	1119	US-10-123-904-352	Sequence 352, App
23	195	4.8	1119	US-10-140-470-352	Sequence 352, App
24	195	4.8	1119	US-09-906-838-294	Sequence 294, App
25	195	4.8	1119	US-09-907-613-294	Sequence 294, App
26	195	4.8	1119	US-09-907-942-294	Sequence 294, App
27	195	4.8	1119	US-10-175-746-352	Sequence 352, App
28	195	4.8	1119	US-10-176-918-352	Sequence 352, App
29	195	4.8	1119	US-10-176-921-352	Sequence 352, App
30	195	4.8	1119	US-10-137-865-352	Sequence 352, App
31	195	4.8	1119	US-10-140-474-352	Sequence 352, App
32	195	4.8	1119	US-09-904-820-294	Sequence 294, App
33	195	4.8	1119	US-09-904-859-294	Sequence 294, App
34	195	4.8	1119	US-09-909-204-294	Sequence 294, App
35	195	4.8	1119	US-10-142-431-352	Sequence 352, App
36	195	4.8	1119	US-10-143-114-352	Sequence 352, App
37	195	4.8	1119	US-09-904-786-294	Sequence 294, App
38	195	4.8	1119	US-09-906-646-294	Sequence 294, App
39	195	4.8	1119	US-09-906-700-294	Sequence 294, App
40	195	4.8	1119	US-10-140-002-352	Sequence 352, App
41	195	4.8	1119	US-09-902-903-294	Sequence 294, App
42	195	4.8	1119	US-09-903-749A-294	Sequence 294, App
43	195	4.8	1119	US-09-903-786-294	Sequence 294, App
44	195	4.8	1119	US-10-142-419-352	Sequence 352, App
45	195	4.8	1119	US-09-902-736-294	Sequence 294, App

## ALIGNMENTS

```
RESULT 1
US-09-884-465A-6
: Sequence 6, Application US/09884465A
: Publication No. US20030077293A1
: GENERAL INFORMATION:
: APPLICANT: Shire Biochem, Inc.
: APPLICANT: Hamel, Josee
: APPLICANT: Brodeur, Bernard
: APPLICANT: Martin, Denis
: APPLICANT: Charland, Nathalie
: APPLICANT: Ouellet, Catherine
: TITLE OF INVENTION: Streptococcus Antigens
: FILE REFERENCE: 055190-0044
: CURRENT APPLICATION NUMBER: US/09/884.465A
: PRIOR FILING DATE: 2001-06-20
: PRIOR APPLICATION NUMBER: 60/212,683
: NUMBER OF SEQ ID NOS: 384
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 6
: LENGTH: 1039
: TYPE: PRT
: ORGANISM: Streptococcus pneumoniae
: US-09-884-465A-6
:
: Query Match
: Best Local Similarity 21.4%: Pred. No. 2.1e-10;
: Matches 185; Conservative 114; Mismatches 283; Indels 283; Gaps 45;
:
: 16 SQTLLICQSRGNGTYPDIKTQSRKMTSNKIRPKRSKTKNTNKGAVGADPPT-----70
: 111: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: 237 SQTLLSYSTASDNN-----TQSVAKGSTS--KPAKSEMLQSLKEL--YDSPSQRVS 285
: 111: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: 71 -DQSFITKRSKTSKTDGIVVDHGHSHFTYADLKG--SFEFLTP---KGASL---121
: 111: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: 286 ESDSLVE-DPAKTIISRPNGVAIPRGGHYHFTYPSKLSALEKTIARNVPLSGSTVSTN 344
: 111: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: 122 AKP-----AAQRAASQGTSK-VADPHNHEFNPADIVADALGYTVRHDDPHNYTL 172
: 111: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: 345 AKPREYVSSIGLSSNSSSLTTSKELSSASDGYIFNPKQIVETATATAYIVRHDDHNYTL 404
: 111: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
: 173 KSSLISGQTQAQKQAVATRLPQTSLSVSTATANGIPGLHPTSPDGFQNGGIVGVTKDSI 232
```

```

Db 405 KSNQIGOPTLPNNSLAT--PSPSLPINPGTSH-----EKHEEGGYGDANRIIAEDSGF 457
Qy 233 LVYDHGHLPIPSFADLROGGMAHVADQYDPAKKAEPATHTOT-----275
Db 458 VMSGIDNNHFFKKDLTE-----EDIKAAQKHELEVKISHNGLDLSSEHDYPG 507
Qy 276 --PELSEREKEYOEKLAYLAELKIDPSTIKRVEYODGKIGLEYPHHDAHVLMDSIEI 333
Db 508 NAKEMDLDKKEIKKIEKIGIMKQYGVKRESI---VVKENKAIITYPHGDH-----553
Qy 334 GKDIIPDHAIEHARELEKHK-VGMDTLRALGDFDEVIIDIVRHDAATPPSPNE---KD 388
Db 554 -----HHADPIDEHKPVG-----IGHSHSYELFKPEEGVAKKE 507
Qy 389 PNMM--KEMLATVYIKLDGSRKDPLOKGLSLPNELETIGITPKIDSPVLOFKKLO 446
Db 588 GNKYVTGEELTNVNL--LKNSTFNQNTFLANGOKRVSPFPPELE-----KKLGI 637
Qy 447 LLMRTGVTDIRFLD-----NMPOLG-----IDISNNLKDISFLS 483
Db 638 NMLVKLITPDGKYLEKVGKVFGEVGNIANFELDOPYLPDQIFKTYTASKDYPVSYDG 697
Qy 484 KYKNLTLVA-----AADNGIEDIRPLGOLP-NLKFLV-----LSNN 518
Db 698 TETVPTSLAYKMASQITFYFPHAGDTYLR-VNPFAPVPGKTDALVRYVDFHGNAYLENN 756
Qy 519 -KISDLS-PLASLHQ-----LOELHIDNNO--ITDLSPVSHKESLTYVDLS 560
Db 757 YKVGEEKLPIPKLNOGTTRTAGNKIPVTFMANAYLDNOSTYIYEV-PILEKENOT-----810
Qy 561 RNADVDLATLQAPKLETLMVNDTRKVSHLDFLKNPN-LSLSINRAOQSL--EGIEASS 617
Db 811 -----DKPSILPQFKRKNKAQENSKLDEKVEPK 838
Qy 618 VIVVEAE-----GNQIKSLVKD-----KQSLTFLDVYG---NOLTSLEGVNNFT 661
Db 839 TSEKVEKEKLESTGNSNSTLEEVPTVDPOEKVAKPAESYKMLEVLFNMDG-----893
Qy 662 ALDILSVSKNOLTVNLSKPKTYNTINIDISHNNISLADKLNGOHIPEALAKNPAYEG 721
Db 894 -----TIELVDSGEV---IKKN--MADE---TGEAPQNGENKPS--EN 928
Qy 722 SMVNGTAEERKAMATKAKESAQEA 746
Db 929 GKYSTGTIVENOPT-ENKPADSLPEA 952

RESULT 2
US-09-884-465A-379
Sequence 379, Application US/09884465A
Publication No. US20030077293A1
GENERAL INFORMATION:
APPLICANT: Shire Biochem, Inc.
APPLICANT: Hamel, Josee
APPLICANT: Brodeur, Bernard
APPLICANT: Martin, Denis
APPLICANT: Charland, Nathalie
APPLICANT: Ouellet, Catherine
TITLE OF INVENTION: Streptococcus Antigens
FILE REFERENCE: 05190-0044
CURRENT APPLICATION NUMBER: US/09/884,465A
CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 60/212,683
NUMBER OF SEQ ID NOS: 384
SOFTWARE: PatentIn version 3.1
SEQ ID NO 379
LENGTH: 1152
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Unknown Organism

```

```

NAME/KEY: MISC_FEATURE
LOCATION: (1)-(1)
OTHER INFORMATION: xaa = Methionine or nothing
NAME/KEY: MISC_FEATURE
LOCATION: (344)-(344)
OTHER INFORMATION: xaa = Glycine or nothing
NAME/KEY: MISC_FEATURE
LOCATION: (345)-(345)
OTHER INFORMATION: xaa = Proline or nothing
US-09-884-465A-379

Query Match
Best Local Similarity 7.2%; Score 292.5; DB 9; Length 1152;
Matches 185; Conservative 114; Mismatches 283; Indels 283; Gaps 45;

Qy 16 SOLTLIACOSRNGNYPTITKOSRKGNKSIKPKKSKTKTKTKGAGVDFPT-----70
Db 350 SOLSYSSIASDNN-----TQSVAKGYS--KPAKNSNLSLKL--YDSPAOKYS 398
Qy 71 -DDGFLITKDSKILSKTDGIVVDHDSHFTFYADLKG--SPFEYLIP---KGASL---121
Db 399 ESDGLVF-DPAKISTRTNGVAILPHODHNFIPYKLSLAEKIRAWPISGTSVSTN 457
Qy 122 AKP-----AAVQRAASQGTSK-VADPHHHYEPNADIVAEALGTYVRHDFHYIL 172
Db 458 AKPNEVYVSSIGSLSSNPSSLTTSKELSSASDGYIFNPKDIVEETATYIVRHGHFHYIP 517
Qy 173 KSSLGTOQAQAQVATRLPOTSSLVSTATANGIRPLHPTSDGFOFGGIVGYTKDSI 232
Db 518 KSNQIGOPTLPNNSLAT--PSPSLPINPGTSH-----EKHEEDGYGDANRIIAEDSGF 570
Qy 233 LVYDHGHLPIPSFADLROGGMAHVADQYDPAKKAEPATHTOT-----275
Db 571 VMSGIDNNHFFKKDLTE-----EDIKAAQKHELEVKISHNGLDLSSEHDYPG 620
Qy 276 --PELSEREKEYOEKLAYLAELKIDPSTIKRVEYODGKIGLEYPHHDAHVLMDSIEI 333
Db 621 NAKEMDLDKKEIKKIEKLAGIMKQYGVKRESI---VVKENKAIITYPHGDH-----666
Qy 334 GKDIIPDHAIEHARELEKHK-VGMDTLRALGDFDEVIIDIVRHDAATPPSPNE---KD 388
Db 667 -----HHADPIDEHKPVG-----IGHSHSYELFKPEEGVAKKE 700
Qy 389 PNMM--KEMLATVYIKLDGSRKDPLOKGLSLPNELETIGITPKIDSPVLOFKKLO 446
Db 701 GNKYVTGEELTNVNL--LKNSTFNQNTFLANGOKRVSPFPPELE-----KKLGI 750
Qy 447 LLMRTGVTDIRFLD-----NMPOLG-----IDISNNLKDISFLS 483
Db 751 NMLVKLITPDGKYLEKVGKVFGEVGNIANFELDOPYLPDQIFKTYTASKDYPVSYDG 810
Qy 484 KYKNLTLVA-----AADNGIEDIRPLGOLP-NLKFLV-----LSNN 518
Db 811 TETVPTSLAYKMASQITFYFPHAGDTYLR-VNPFAPVPGKTDALVRYVDFHGNAYLENN 869
Qy 519 -KISDLS-PLASLHQ-----LOELHIDNNO--ITDLSPVSHKESLTYVDLS 560
Db 870 YKVGEEKLPIPKLNOGTTRTAGNKIPVTFMANAYLDNOSTYIYEV-PILEKENOT-----923
Qy 561 RNADVDLATLQAPKLETLMVNDTRKVSHLDFLKNPN-LSLSINRAOQSL--EGIEASS 617
Db 924 -----DKPSILPQFKRKNKAQENSKLDEKVEPK 951
Qy 618 VIVVEAE-----GNQIKSLVKD-----KQSLTFLDVYG---NOLTSLEGVNNFT 661
Db 952 TSEKVEKEKLESTGNSNSTLEEVPTVDPOEKVAKPAESYKMLEVLFNMDG-----1006
Qy 662 ALDILSVSKNOLTVNLSKPKTYNTINIDISHNNISLADKLNGOHIPEALAKNPAYEG 721
Db 1007 -----TIELVDSGEV---IKKN--MADE---TGEAPQNGENKPS--EN 1041
Qy 722 SMVNGTAEERKAMATKAKESAQEA 746

```

Db 1042 GKVSTGTVENOPT-ENKPADSLPEA 1065

## RESULT 3

US-09-884-465A-381  
Sequence 381, Application US/09884465A  
Publication No. US2003007293A1  
GENERAL INFORMATION:  
APPLICANT: Shire Biochem, Inc.  
APPLICANT: Hamel, Josee  
APPLICANT: Brodeur, Bernard  
APPLICANT: Martin, Denis  
APPLICANT: Charland, Nathalie  
APPLICANT: Ouellet, Catherine  
TITLE OF INVENTION: Streptococcus Antigens  
FILE REFERENCE: 055190-0044  
CURRENT APPLICATION NUMBER: US/09/884,465A  
PRIOR FILING DATE: 2001-06-20  
PRIOR APPLICATION NUMBER: 60/212,683  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 384  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 381  
LENGTH: 1238  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Unknown Organism  
NAME/KEY: MISC\_FEATURE  
LOCATION: (1)..(1)  
OTHER INFORMATION: Xaa = Methionine or nothing  
NAME/KEY: MISC\_FEATURE  
LOCATION: (430)..(430)  
OTHER INFORMATION: Xaa = Glycine or nothing  
NAME/KEY: MISC\_FEATURE  
LOCATION: (431)..(431)  
OTHER INFORMATION: Xaa = Proline or nothing  
US-09-884-465A-381

Query Match 7.2%; Score 292.5; DB 9; Length 1238;  
Best Local Similarity 21.4%; Pred. No. 2.7e-10;  
Matches 185; Conservative 114; Mismatches 283; Indels 283; Gaps 45;

QY 16 SOLTLLACOSRNGNTYPIKTKOSRKGMTSNKIKPIKSKTKTKTKHGAGVDEPT----- 70  
DB 436 SOLTSSVSTASDNN-----TOSVANGSTS--KPAKSKSENLQSLKEL--YDSPAQRYS 484  
QY 71 -DDGFIITKDSKILSKTDGIVVDHGHSHFIFYADLKG--SPEFYLP--KGSJL--- 121  
DB 485 ESDGLVF-DPAKIIISRTGPAVAIPHGDIHFIPIYKLSLAEKIKRMPISGTSVSTN 543  
QY 122 AKP-----AAQRAASQGTSK-VADPHHHYEPNPADIVAEALGYTVRHHDFHYIL 172  
DB 544 AKPNEVSSISGLSSNPSSILTSKELSSADGIFNPDIYEETATYIVRHGDHFHYIP 603  
QY 173 KSSLGQTOQAOKOYATRLPOTSSLSVSTATANGIPGLHFPPISDGQFGOGGIVGTAKSI 232  
DB 604 KSNQIGQPTLPNNLSLAT--PSPSLPINTGTSI---EKHEEDGFGFANRLIADESGF 656  
QY 233 LVVDHGHLPISFADLRGGMHNAVADQYDPAKKAERPAETHOT----- 275  
DB 657 VMSHGDHNNHYFFKKDLTE-----EQIKAAQKHLEEVKTSINGLSDLSSEHODYPG 706  
QY 276 --PELSEKEKEQEKALYLAELKGLDPESTIKRVEYODGKGLGLEYPHHDHAYMLSDIEI 333  
DB 707 NAKEMKDLDKLEEKIADIMQYKRSI---VYKKEKNAIIFYPHGDH----- 752  
QY 334 GKDIIDPNAIEHARELEKHK-VGMDTLRALGFDDEVILDIYRTHAPRPPPSNE---KD 388  
DB 753 -----HHADPIDEHKPYG-----IGHSHSNVELEKPEEGVAKKE 786  
QY 389 PNM--KEMLATVILKLGSRKDLQKGLSLPNEFLTGIGFTPIPKDISVLOFKKILQ 446

Db 787 GUKVYTGEEITNVNL-----LKNSTFNNOFTLANQOKRVSEFPPEL-----KIKGI 836

QY 447 LAMTKGVTDYRFLD-----NMPQLEG-----IDISQNNLKDIFSL 483  
DB 837 NMLVLTIPEDGKVLKESKVGEGVGNLANFELDPYLPQGTFFYTTIASDYEPVSDG 896  
QY 484 KYKNLTLVA-----AANGIEDIRPLGQLP-NLKELY-----LSNN 518  
DB 897 TETVPTSLAYKMASQTIFFPHAGDTYLR-VNPOPAVRKGTDALRVDFEFGNAYLENN 955  
QY 519 -KISDLS-PLASLHQ-----LOELHIDNNO--ITDLSPVSHKESITLVYDLS 560  
DB 956 YVGEIKLPIPKLNGTTRTAGNKIPVTFMANAYIDNSTIVEX-PLKEENOT----- 1009  
QY 561 RNADVADLTLPKLETLMVNDTKVSHLDFLKNPN-LSSLSINRAQLQSL--EGTEASS 617  
DB 1010 -----DKPSILPQKRNKAQNSKLDEKVEPK 1037  
QY 618 VIVRYEAE-----GNQIKSLVKD-----KQSLTFLDVYG-----NQLTSLGCVNNFT 661  
DB 1038 TSEKYEKELSTGNSSTSLVEEPTVDVQEKYAKFAESYGKLEVLNMDG----- 1092  
QY 662 ALDILSVSKNQLTNVLSKPKNTYTNIDISHNNTSLADKLNEDHIEPAIAKNFPAYEG 721  
DB 1093 -----TEIYLPGEV-----IKKN-----MADF-----TGEARQNGENKPS--EN 1127  
QY 722 SMVNGTAEKKAAMATKAKESAQEA 746  
DB 1128 GKVSTGTVENOPT-ENKPADSLPEA 1151

## RESULT 4

US-09-884-465A-382  
Sequence 382, Application US/09884465A  
Publication No. US2003007293A1  
GENERAL INFORMATION:  
APPLICANT: Shire Biochem, Inc.  
APPLICANT: Hamel, Josee  
APPLICANT: Brodeur, Bernard  
APPLICANT: Martin, Denis  
APPLICANT: Charland, Nathalie  
APPLICANT: Ouellet, Catherine  
TITLE OF INVENTION: Streptococcus Antigens  
FILE REFERENCE: 055190-0044  
CURRENT APPLICATION NUMBER: US/09/884,465A  
PRIOR FILING DATE: 2001-06-20  
PRIOR APPLICATION NUMBER: 60/212,683  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 384  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 382  
LENGTH: 1365  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Unknown Organism  
NAME/KEY: MISC\_FEATURE  
LOCATION: (1)..(1)  
OTHER INFORMATION: Xaa = Methionine or nothing  
NAME/KEY: MISC\_FEATURE  
LOCATION: (557)..(557)  
OTHER INFORMATION: Xaa = Glycine or nothing  
NAME/KEY: MISC\_FEATURE  
LOCATION: (558)..(558)  
OTHER INFORMATION: Xaa = Proline or nothing  
US-09-884-465A-382

Query Match 7.2%; Score 292.5; DB 9; Length 1365;  
Best Local Similarity 21.4%; Pred. No. 3e-10;  
Matches 185; Conservative 114; Mismatches 283; Indels 283; Gaps 45;

QY 16 SOLTLLACOSRNGNTYPIKTKOSRKGMTSNKIKPIKSKTKTKTKHGAGVDEPT----- 70  
DB 436 SOLTSSVSTASDNN-----TOSVANGSTS--KPAKSKSENLQSLKEL--YDSPAQRYS 484

```

Db 563 SOLSYSSTASDNM-----TQSAKSTG---KPAKSENLSLKL--YDSPAQRYS 611
Qy 71 -DDGFILTKDSKILSKTDGIVVDHSHFIFYADLK--SPEYILP---KGASL--- 121
Db 612 ESDGLVF-DPAKILSRTPNGVAIPHGDIHFIPYSKLSALEEKIARVVISGGSIVSTN 670
Qy 122 AKP-----AVNORAASQTSK--VADPHNHYEFPADIVAEADALGYVRHDDHFHYIL 172
Db 671 AKPNEVVSLSLSSNSSLTTSKELSSASDGYIFNPKDIYEATATAYVRHGDHFHYIP 730
Qy 173 KSLSGQTOAQAKOVATRLPOTSSLVSTATANGIPGLHPTISDGFQNGGIVGVTKDSI 232
Db 731 KSNIGQPTLPNNSLAT--PSPSLPINPGTSH-----EKHEDEGYGFANRIIAEDES GF 783
Qy 233 LVYHDGHLPIFSADLRGGGMAHVAADQYDPAKKAERAEHTQT----- 275
Db 784 VMSHGDHNYHFFKKDLTE-----EQIKAQKHLEEVKTSHGSLDLSLSSHEODYPG 833
Qy 276 --PELSEKEKYOERKLAELAELGIDPSTIKRVETODGKLEYPHNDHAHVLMSDIEI 333
Db 834 NAKEMKDLDKKIEKTAGIMKOYGVKRESI---VNNKKNAILIYPRGDH----- 879
Qy 334 GKDIIPDPAIEHARELEKHK--VGMOTLRALGFDEEVLIDIVRTHDAPTPPSNE---KD 388
Db 880 -----HHADPIDENKRPVG-----IGHSHSNVELFKPEEGVAKKE 913
Qy 389 PNMM--KEMLATYIKLDGSRKDPLOKRGSLLENLETIGTIPRIDISVLOFFKKLQ 446
Db 914 GNKYVTGELTNVNL--LKNSTFNNQNTFLANGOKRVSFSPPELE-----KKLGI 963
Qy 447 LMTKGTGVTDRFLD-----NMPOLEG-----IDISQNNLKDISFLS 483
Db 964 NMLVYLITPDGKYLEKVGKVFGEVGNIANFELDQYLRPGOTFKYTLASDYEVSYDG 1023
Qy 484 KYKNULTVA-----AADNGIEDIRPLGOLP-NLKFVY-----LSNN 518
Db 1024 TETVPTSLAYKMASOTIFEPFHAGDTYLR--VNPQFAVPRKGTDALRVFDEFHGNAYLENN 1082
Qy 519 -KISLDS-PLASLHQ-----LOELHIDNNQ--ITDLSVSHKESLTVVDLS 560
Db 1083 YKVGELKLPKPLNOGTTTTRAGNKIPVTFMANAYLIDNOSTIYEV-PILEKNOT----- 1136
Qy 561 RNADVDLATLQAPKLETLVNDTKVSHDLFLKNPN-LSSLISINRAQLOSL--EGIEASS 617
Db 1137 -----DKPSILPQFKRNKAQNSKLDKVEEPRK 1164
Qy 618 VIVRVEAE-----GNOISLVYKD-----KOSLTFPLDTG---NQLTSLEGVNFT 661
Db 1165 TSEKVEKEKLESETGNSSTSLTEVPYDVOEKVAKFAESYGMKLENVLFNMDG----- 1219
Qy 662 ALDILSVSKNQLTNVNLKPKNTVTNIDISHNNISLADLKINEOHIPALAKNFPVAYEG 721
Db 1220 -----TIELYIPSGEV---IKKN---MADE---TGAPRGNGENKRS--EN 1254
Qy 722 SMVNGTAEKAKAMATKAKESAQEA 746
Db 1255 GKYSTGTVENQPT-ENKPADSLPEA 1278

```

```

RESULT 5
US-09-884-465A-378
Sequence 378, Application US/09884465A
Publication No. US20030077293A1
GENERAL INFORMATION:
APPLICANT: Shire Biochem, Inc.
APPLICANT: Hamel, Josee
APPLICANT: Brodeur, Bernard
APPLICANT: Martin, Denis
APPLICANT: Charland, Nathalie
APPLICANT: Ouellet, Catherine
TITLE OF INVENTION: Streptococcus Antigens
FILE REFERENCE: 053190-0044
CURRENT APPLICATION NUMBER: US/09/884,465A

```

```

CURRENT FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: 60/212,683
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 384
SOFTWARE: Patencin version 3.1
SEQ ID NO 378
LENGTH: 1378
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Unknown Organism
NAME/KEY: MISC_FEATURE
LOCATION: (1)..(1)
OTHER INFORMATION: Xaa = Methionine or nothing
NAME/KEY: MISC_FEATURE
LOCATION: (570)..(570)
OTHER INFORMATION: Xaa = Glycine or nothing
NAME/KEY: MISC_FEATURE
LOCATION: (571)..(571)
OTHER INFORMATION: Xaa = proline or nothing
US-09-884-465A-378
Query Match
Best Local Similarity 21.4%; Pred. No. 3,1e-10;
Matches 185; Conservative 114; Mismatches 283; Indels 283; Gaps 45;
Qy 16 SOLTLIACSGRNGTYPIKTKOSKQMTSNKIKPIKSKTKTKTKHGVAGVDEPPT----- 70
Db 576 SOLSYSSTASDNM-----TQSAKSTG---KPAKSENLSLKL--YDSPAQRYS 624
Qy 71 -DDGFILTKDSKILSKTDGIVVDHSHFIFYADLK--SPEYILP---KGASL--- 121
Db 625 ESDGLVF-DPAKILSRTPNGVAIPHGDIHFIPYSKLSALEEKIARVVISGGSIVSTN 683
Qy 122 AKP-----AVNORAASQTSK--VADPHNHYEFPADIVAEADALGYVRHDDHFHYIL 172
Db 684 AKPNEVVSLSLSSNSSLTTSKELSSASDGYIFNPKDIYEATATAYVRHGDHFHYIP 743
Qy 173 KSLSGQTOAQAKOVATRLPOTSSLVSTATANGIPGLHPTISDGFQNGGIVGVTKDSI 232
Db 744 KSNIGQPTLPNNSLAT--PSPSLPINPGTSH-----EKHEDEGYGFANRIIAEDES GF 796
Qy 233 LVYHDGHLPIFSADLRGGGMAHVAADQYDPAKKAERAEHTQT----- 275
Db 797 VMSHGDHNYHFFKKDLTE-----EQIKAQKHLEEVKTSHGSLDLSLSSHEODYPG 846
Qy 276 --PELSEKEKYOERKLAELAELGIDPSTIKRVETODGKLEGYEVRHNDHAHVLMSDIEI 333
Db 847 NAKEMKDLDKKIEKTAGIMKOYGVKRESI---VNNKKNAILIYPRGDH----- 892
Qy 334 GKDIIPDPAIEHARELEKHK--VGMOTLRALGFDEEVLIDIVRTHDAPTPPSNE---KD 388
Db 893 -----HHADPIDENKRPVG-----IGHSHSNVELFKPEEGVAKKE 926
Qy 389 PNMM--KEMLATYIKLDGSRKDPLOKRGSLLENLETIGTIPRIDISVLOFFKKLQ 446
Db 927 GNKYVTGELTNVNL--LKNSTFNNQNTFLANGOKRVSFSPPELE-----KKLGI 976
Qy 447 LMTKGTGVTDRFLD-----NMPOLEG-----IDISQNNLKDISFLS 483
Db 977 NMLVYLITPDGKYLEKVGKVFGEVGNIANFELDQYLRPGOTFKYTLASDYEVSYDG 1036
Qy 484 KYKNULTVA-----AADNGIEDIRPLGOLP-NLKFVY-----LSNN 518
Db 1037 TETVPTSLAYKMASOTIFEPFHAGDTYLR--VNPQFAVPRKGTDALRVFDEFHGNAYLENN 1095
Qy 519 -KISLDS-PLASLHQ-----LOELHIDNNQ--ITDLSVSHKESLTVVDLS 560
Db 1096 YKVGELKLPKPLNOGTTTTRAGNKIPVTFMANAYLIDNOSTIYEV-PILEKNOT----- 1149
Qy 561 RNADVDLATLQAPKLETLVNDTKVSHDLFLKNPN-LSSLISINRAQLOSL--EGIEASS 617
Db 1150 -----DKPSILPQFKRNKAQNSKLDKVEEPRK 1177

```







202 TANGIGLHPPS-----DGFQNGQIVGVTKDLSLVDDHGHLPISFADLRQGGMAH 255  
224 TPSPSLPINSISHEKHEDEGYGFDFANRIAEDESGFVMSHGHNHYFFPKDLTE----- 278  
256 VADQYPAKAEKPAETHOT-----PELSEKKEOEKLAIAEKL 296  
279 -----EIKAKQCHLEBKVSHNGLDLSLSEHEDYPSNAKEMKDLCKLEEKTAGIMKOY 333  
297 GIDPSTIKRVETDGLGLEYPHNDHANVLMLSDEIGKIDPPHAIEHARELEKHK-VG 355  
334 GVKRESI-----VVKKEKNALITPHGDH-----HHADPIDEKHPYG 369  
356 MDLIRALGDEEYIIVTRHDAPTPPESNE-----KDPNMM-KEKLAIVIKIDLSKRD 409  
370 -----IGHSHNYELFKPEEGYAKKEGKVYTGEBELTVVNL-----LKNS 410  
410 PLQKGSLLPNETLGIGTTPKIDISPVLOFKKLQMLMTKTGVDFPLD----- 461  
411 TTNQNFITLANGOKRVSPFPPEL-----KKLGIMLVKLITPDGKYLEKVGSKVG 463  
462 -----NMPQLEG-----IDISQNLKDISFLSKYKNLTLVA-----A 493  
464 EGVGNIANFELDQPYLPGCFKTYTISKDYPEVSDGTFTVPTSLAYKMASQTIFFPFA 523  
494 ADNGIEDIRPLGOLP-NLKFVY-----LSNN-KISDLS-PLASIHQ----- 531  
524 GDYTLR-VNQFAVPKGTALVRFDEPHGNAYLENNYKGEIKLIPKLNQGTTPAGN 582  
532 -----LOELHIDNQG--ITDLSPVSHKESLTVVDSLRNADVLTALQAPKLETLAVNDT 583  
583 KIVYTFMANVYLDNOSTYIEV-PILEKENOT----- 613  
584 KVSHDLFLKNPN-LSSLINRAQLOSL--EGIEASSVIVREAE-----GNOIKSLVJK 635  
614 -----DKPSILPOFKRNKAQENSKLDEKVEEPTSEKVEKELSENGNSTSTLE 664  
636 D-----KGSITFLDYVG-----NOLISLEGVNNFTALDILSVSKNQLTNVLSKPNKT 684  
665 EVPTVDPVQKRVAKFAESYGMKLENLVFNMDG-----TTELYPSPSE 706  
685 VTNIDISHNNISLADLKLNQOIHPEAIKKNFPAVYEGSMVNGTAEKAMATKAKESAQ 744  
707 V-----IKKN-----MADF-----TGEAPQNGEKPKPS--ENGKYSTGTVENOPT-ENKPADSLP 753  
QY 745 EA 746  
Db 754 EA 755

RESULT 10  
US-09-884-465A-8  
; Sequence 8, Application US/09884465A  
; Publication No. US2003007293A1  
; GENERAL INFORMATION:  
; APPLICANT: Shire Biochem, Inc.  
; APPLICANT: Hamel, Josee  
; APPLICANT: Brodeur, Bernard  
; APPLICANT: Martin, Denis  
; APPLICANT: Charliet, Nathalie  
; APPLICANT: Ouellet, Catherine  
; TITLE OF INVENTION: Streptococcus Antigens  
; FILE REFERENCE: 055190-0044  
; CURRENT APPLICATION NUMBER: US/09/884,465A  
; CURRENT FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: 60/212,683  
; PRIOR FILING DATE: 2000-06-20  
; NUMBER OF SEQ ID NOS: 384  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8  
; LENGTH: 838  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-884-465A-8

Query Match 5.5%; Score 224; DB 9; Length 838;  
Best Local Similarity 20.3%; Pred. No. 3.5e-06;  
Matches 173; Conservative 122; Mismatches 321; Indels 238; Gaps 40;  
41 GMTSKIKIPKSKSKTKNTKHKGVAGVDEPTDDGFLITLDSKILSTKDOGIYVDDHDSHF 100  
Db 48 GOKAENLTPDEVSKREG-----INAEQIVIKITDGGYVTSIGDHYH 89  
QY 101 IFYADLKSPFEYLIPKASLAKPAVAQAASQGTSKVADHNNHYEFPADIVADALGY 160  
Db 90 -----YNGKPYALITSEEL-----MKDP--NYOLKSDIVAEIKGY 126  
QY 161 TVRHDDHFHYILKSSLSGQTOAQAKOVATRLPQTS-----SLVSTATANGIPGLHFP 212  
Db 127 VIKYDGKYYVYLKDAHANDNIRTKKEIKRQKQESHNNHNSRADNAVAARAQO-----RYT 182  
QY 213 TSDGFQFQNGQIVGVTKOSILVDHGHLPISFADLRQGGMAHADVQDPKAKKEPAET 272  
Db 183 YDDGYIFNASDIIEDTGDYIVPHGDHYIIPKNELSASELA-AAEAAYWNGKQSRPSS 241  
QY 273 HQ-----TPELSER-----EKEYOEKLAIAEKLIDPSTIKRVETQDCKL----- 313  
Db 242 SSYMANPVQPRLESHNNLVTPTTHQNGENISSLRRELYAKPLSERIVES-DGLITDPA 300  
QY 314 -----GLEYPHDHAHVL--MLSDIE--IGKIDP-----DPHAIEHAR-ELEKHKVG 355  
Db 301 QITSTRAGVAVPHGNHVFIEYEQMSELEKRIARITLRYRSHWVPDSRPOSPQST 360  
QY 356 MDLIRALGDEEYIIVTRHDAPTPPESNEKDPNMMKEMLATYIKIDLSGRKRPLOKKG 415  
Db 361 PEPSPSL-----QAPNPQAPSNPTDEKIVE--AAKQVDGY--VFEENG 403  
QY 416 LLLPNETLGIGTTPKIDISPV-----LOFKLKO-LMTKTGV--TD----- 456  
Db 404 VS-----RTPADLSAETAAGIDSKLAKESLSHKLGAKKTDLPSSDREFYKA 453  
QY 457 YFELDNMQLEGIDISQNLKDISFLSKYKNLTLVAADNGIEDIRPLGOLP-NLKFVLS 516  
Db 454 YDLARHQ-----DLIDNKGROVDF-----EVIDNLLERLKD-----VS 488  
QY 517 ANKI-----SDLSPLASLMOLOELHIDNNOITDLSPVSHKESLTVVDSLRNADVLTALQ 572  
Db 489 SDKVKLVDDILAFIAPIRHPERLGRPNQOITYTD-----DEIQAKLAGKTYTDEGYIED 543  
QY 573 PKEITLWVNDTKV-----SHLDPLKNPNISLISINRAQLOSL--GIEASSVIVREAE 627  
Db 544 PRDITSDEGDATVTPHMTSHHWIKD-SLSEAEKRAAQAAYAKKGLTPPST----- 593  
QY 628 QIKSLVLKDKQSLTFLDYVGNQLTSLEGV-----NFTALDILSVK--NQLTNVLSK 680  
Db 594 -----DHQDS-----GN--TEAKGAFAIYRVAKAKVPLDRMPYNIQYTVYEKN 636  
QY 681 PKTYTINIDISHN-----NISLADLKLNQOIHPEAIKKNFPAVYEGSMVNG 725  
Db 637 GSLITPHDYHNIKFEWDEGLYAPKGYSLIEDLATVATKYVE-----HNERPHSNG 691  
QY 726 NQTAEEKAMATKAKESAQESHDYNNHNTY--EDEGHAHEHRKRD-----HDHEH 778  
Db 692 FQNASDHWKKNKADQSKPDEKHEDEVSEPTHPESDKENHAGINPABADNLYKPSDTE 751  
QY 779 EDENEAKDEQNHAD 792  
Db 752 ETEEEAEDTTDEAE 765

RESULT 11  
US-09-765-272-56  
; Sequence 56, Application US/09765272  
; Patent No. US20020061545A1  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
US-09-765-272-56





```

QY 605 AQLSLEIEASSVIVAEAGNOIKSLVKDKGSLTFELDTGNOITSL-EGVNNFTAL 663
DB 1050 AQLSSIPG-----ELITKLSFLEKLELNNNNLRLRPEQISKLTKL 1089
QY 664 DILSSKNOITNV--NLSKPKVTYNTIDISHNNI-----SLADKLNEQHPEIAKNFP 716
DB 1090 VFLSVARNKLEYIPELSQ-LKSLTLDLHNNIRDFVDGMEHLEISLNT----- 1139
QY 717 AVEGSMVNGTAEKKAAMATKAKESQAESESHDYNNHNTY 758
DB 1140 ---SSNARGNSSLE-----NSFYHNMXY 1159

RESULT 14
US-09-905-291A-294
; Sequence 294, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertschen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02

```

```

; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO: 294
; LENGTH: 1119
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-291A-294

Query Match          4.8%; Score 195; DB 9; Length 1119;
Best local similarity 21.6%; Pred. No. 0.00036;
Matches 105; Conservative 90; Mismatches 149; Indels 142; Gaps 23;

QY 286 QEKATVLAELKIGIDPSTIKRVEITQDGKLGLEYPHHDHAYLMLSDIEIGKIDPPPAIEH 345
DB 62 RKRLARLEPL--PSVVARLDLS-----HNRSLFKASSMS-----HLOS 99
QY 346 ARELEKHKVGMOTLRALGFDEEVILIVRTHDAPTEPPSNEKDPNMKKEWLATVIKLDG 405
DB 100 LREVKLNNELETIPNG-----PVSAN-----ITLSEL-AG 130
QY 406 SRKDPLOKRGISLPLNETIGCTPIPKDISPVLOEFKKLQOLMTKGYVDYR--FLDN- 462
DB 131 NRIVEILPEHLKEPQSLSETLDLSSNNISELOTAPALQLYLYLNSNRYTSMERGFDNL 190
QY 463 -----MPOLEGIDISQNNIKDISFLS-----KYKNLTVAAA 494
DB 191 ANTLIVLKLNNRISALPRPKFKLPQLOHLELNKKNVDFLFGSLGALKSLKM--Q 247
QY 495 DNGIEDIR--PLGQLPNIKFLVLSNNKISDLS--LASLHQDELHIDNNQITDLSVSH 550
DB 248 RNVGTKLMDGAFMGLSNMELIQLDHNHNLTEYKGLMGLMLDELHLSQNAIRISPDW 307
QY 551 K--ESLTVVDLSRN-----ADVDLATLQAPKLETLVMDTVSH-----LDF 590
DB 308 EFCQSLSELDLTFHNLRLD--DSSFGLSLNLTHTHIGNKRYIADCAFGLSLKTLD- 365
QY 591 LKNPNLSLSLNRQQLSLEIEASSVIVAEAGNOIKSLVKDKG--SLTFELDTG 648
DB 366 LKNNDISWTIEDMGATSGDLKLR-----RLIQGRIRISITKKAFTGLDALEHLDLSD 419
QY 649 NQLTSLEGVNFTALDILSVSKNQLTNVNSKPKVTYNTIDISHNNISLADKLNEQHIF 708
DB 420 NAIMSLQG-NAFS-----QMKKLQDLHLNNTS-----LLDCQGL--KWL 456
QY 709 EAIKRN 714
DB 457 QWVAEN 462

RESULT 15
US-09-902-853-294
; Sequence 294, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter

```

```

APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavon, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/902,853
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US/09/665,350
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 294
LENGTH: 1119
TYPE: PRT
ORGANISM: Homo Sapien
US-09-902-853-294

```

Query Match 4.8%; Score 195; DB 9; Length 1119;  
 Best Local Similarity 21.6%; Pred. No. 0.00036;  
 Matches 105; Conservative 90; Mismatches 149; Indels 142; Gaps 23;

```

QY 286 QEKAVYAEKGLDPSIKRVEODGKLGIEYHHHAHVMLSIEIGKIDPDPAIEH 345
Db 62 RKRRLARPEPL--PSVVARLDLS-----HNRLSFKASMS-----HLQS 99
QY 346 ARELEKHKVGMDFLRALGFDEEVLIDIVRTHDAPTFPPSNKEDPNMMEKMLATVIKLDLG 405
Db 100 LREVKLNNELETTPNLG-----PVSAN-----ITLLSL-AG 130

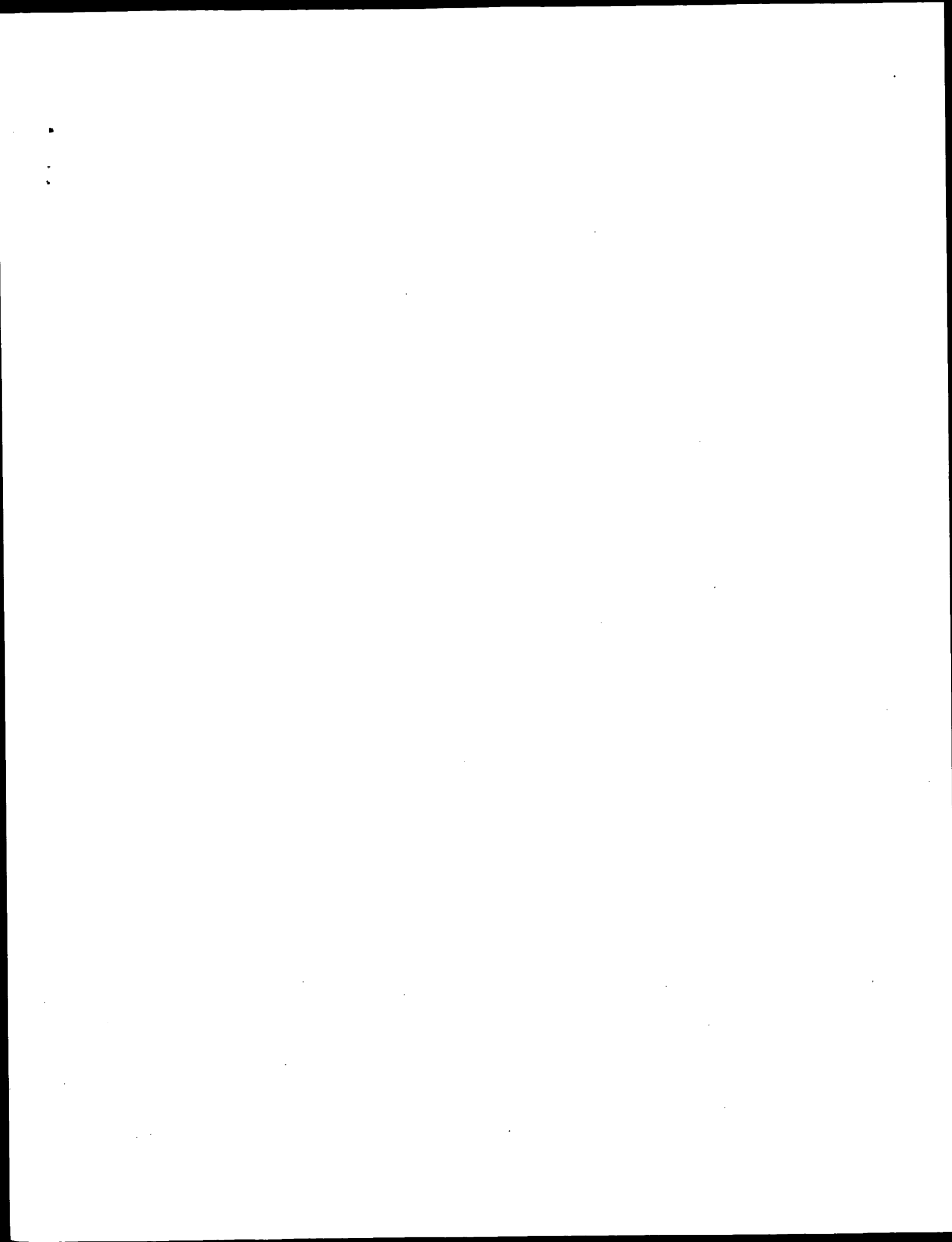
```

```

QY 406 SRKDPLOKRGSLPLNPIETIGTFTPIKDISPVLOEFKKQLMLTKGTVDYR--FLDN- 462
Db 131 NRIVEILLPEHLKEPQSLSETIDLSNNITSELOTAPFALQLYKLYNSRVTSMPEGYFDNL 190
QY 463 -----MPOLEGIDISNNLKDISFIS-----KYKULTVAAA 494
Db 191 ANTLVLKLNKRNRISALTPPKMKFLPOLQHLNKNKIKNVGDLTFQGLGALKSLKM--O 247
QY 495 DNGIEDIR--PLGOLPNIKFLVLSNNKISDLSP--IASLHLOELHIDNNQIDLSPVSH 550
Db 248 RNVGTRKLMGCAFVGLSNMHEILQLDHNNLTETTKGWLIVGLIMDELHLSQNAIRISPDAM 307
QY 551 K--ESLTVVDLSRN-----ADVDLATIQAPKLEFLMNVDTVSH-----LDE 590
Db 308 EFCCKLSLDELITRNHLSRLD--DSSFGLSLNLTLHIGNNVSYIADCAFGLSSLEKTLDP- 365
QY 591 LKNNPNLSLSINRAQLOSLEGIEASSVIVRAEAGNOIKSLVAKDKOG--SLTFPDVYG 648
Db 366 LKNEISWTIEDMNGAFSGDKLR-----RLLOGNRIRISITRKAFITGLDALEHLDLSO 419
QY 649 NQLTSEGVNNFTALDILSVSKNQLTNNVLSKPNKTVTNIDISHNNISLADLKLNQOHIP 708
Db 420 NAIMSLQG-NAFS-----QMKKLOQLHLNTSS-----LLCCQQL--KWLP 456
QY 709 EATAKN 714
Db 457 QWVAEN 462

```

Search completed: May 19, 2003, 08:24:04  
 Job time : 239.541 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: May 19, 2003, 07:51:03 ; Search time 12.7448 Seconds

(without alignments)  
1883.841 Million cell updates/sec

Title: US-09-645-835A-7

Perfect score: 4265  
Sequence: 1 MKINKRYLKGSAALILSYC.....KLALLKGSNPSVSKEKIN 816

Scoring table: BLOSUM62  
Gap 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfillseq1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4163	97.6	796	4	US-08-961-083-56
2	2661.5	62.4	763	4	US-08-961-083-66
3	1186.5	27.8	447	4	US-08-961-083-182
4	162	3.8	3696	4	US-09-134-001C-5080
5	160	3.8	10182	4	US-09-134-001C-3159
6	152	3.6	783	6	5231168-2
7	147	3.4	812	4	US-09-513-783A-6
8	145.5	3.4	1073	4	US-09-541-782-6
9	145.5	3.4	1073	4	US-09-723-820-6
10	144	3.4	1040	4	US-08-961-083-118
11	142.5	3.3	2314	4	US-09-816-703A-2
12	140	3.3	1503	4	US-08-976-255-14
13	140	3.3	2431	1	US-07-920-281C-2
14	140	3.3	2431	1	US-08-466-277-2
15	137.5	3.2	1610	4	US-09-513-783A-22
16	136	3.2	812	4	US-09-513-783A-4
17	136	3.2	1056	4	US-09-513-783A-32
18	135.5	3.2	740	1	US-08-257-073-5
19	135.5	3.2	1183	2	US-08-447-031A-2
20	135	3.2	2308	2	US-08-015-973-1
21	135	3.2	2308	2	US-08-448-164-1
22	135	3.2	2308	4	US-08-081-929-2
23	133.5	3.1	1507	6	5268270-2
24	133.5	3.1	1664	1	US-09-599-652-2
25	133.5	3.1	1664	2	US-08-642-846-2
26	133.5	3.1	1664	4	US-09-264-604-2
27	133	3.1	1093	5	PCT-US93-03077-1

28	132	3.1	1164	4	US-09-457-708-2	Sequence 2, Appl
29	132	3.1	1338	1	US-08-471-033-50	Sequence 50, Appl
30	132	3.1	1338	2	US-08-471-044-50	Sequence 50, Appl
31	132	3.1	1338	2	US-08-463-483A-50	Sequence 50, Appl
32	132	3.1	1338	2	US-08-470-046A-50	Sequence 50, Appl
33	132	3.1	1338	2	US-08-470-566B-50	Sequence 50, Appl
34	132	3.1	1338	2	US-08-469-334-50	Sequence 50, Appl
35	132	3.1	1338	3	US-09-300-528-50	Sequence 50, Appl
36	131.5	3.1	2032	4	US-09-071-035-458	Sequence 468, App
37	131.5	3.1	2032	4	US-09-071-035-462	Sequence 468, App
38	131.5	3.1	2032	4	US-09-071-035-466	Sequence 468, App
39	130.5	3.1	1435	2	US-08-568-459A-4	Sequence 4, Appl
40	130.5	3.1	1435	2	US-08-487-826B-4	Sequence 4, Appl
41	130.5	3.1	1435	4	US-09-210-288-4	Sequence 4, Appl
42	130	3.0	2115	4	US-09-324-867-5	Sequence 4, Appl
43	129	3.0	688	3	US-09-141-047-8	Sequence 8, Appl
44	129	3.0	2660	2	US-08-826-267-2	Sequence 2, Appl
45	128.5	3.0	916	1	US-08-188-228-48	Sequence 48, Appl

#### ALIGNMENTS

```
RESULT 1
US-08-961-083-56
; Sequence 56, Application US/08961083
; Patent No. 6159469
;
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 796 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-083-56

Query Match 97.6%, Score 4163, DB 4, Length 796;
Best Local Similarity 99.9%, Pred. No. 2.2e-305;
Matches 795; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 21 STEGLYARFYKKNRYSYIDGKATOKTENLTPDEVSREKREGINARDIYIKITDQGYVT 80
DB 1 STEGLYARFYKKNRYSYIDGKATOKTENLTPDEVSREKREGINARDIYIKITDQGYVT 60
```





Db 646 GYLEDLATAVKKYVHEPHNPHSDNGFNASDHVQNRKNGQADTQTEKPESEKQTER 705  
 QY 729 PVEET-----PAEPVPOVETKEVPAOLKEAFLAKATD 763  
 Db 706 PEETPEEKPEKPESEKPEKTEPESEPESEPEVETKEVKEKLEADLGLKIOD 763

## RESULT 3

US-08-961-083-182  
 ; Sequence 182, Application US/08961083  
 ; Patent No. 6159469  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Choi et. al.  
 ; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
 ; NUMBER OF SEQUENCES: 452  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
 ; COMPUTER: HP Vectra 486/33  
 ; OPERATING SYSTEM: MSDOS version 6.2  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/961,083  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Brookes, A. Anders  
 ; REGISTRATION NUMBER: 36,373  
 ; REFERENCE/DOCKET NUMBER: PB340P2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (301) 309-8504  
 ; TELEFAX: (301) 309-8512  
 ; INFORMATION FOR SEQ ID NO: 182:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 447 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-961-083-182

Query Match 27.8%; Score 1186.5; DB 4; Length 447;  
 Best Local Similarity 41.0%; Pred. No. 2.1e-81;  
 Matches 263; Conservative 62; Mismatches 110; Indels 207; Gaps 11;

QY 24 LGYQATVKNRNVSYIDKQATQKTELTPEVSKREGINAEOIYIKITDGYVTSHG 83  
 Db 1 LNHRSGENKNNRVSIVDSQSOKSENLTPOVSKKEIOAEQIYIKITDGYVTSHG 60  
 QY 84 DHHYVNGKPYDAIIEELLKMDPNYKLDKEDIVNEVGKGYIKYDKYVYLKDAHA 143  
 Db 61 DHHYVNGKPYDALFSEELLKMDPNYKLDKADIVNEVGKGYIKYDKYVYLKDAHA 120  
 QY 144 DNVTKEIRNKOESHOHREGTPRNDGAVVALARSQRTYTDGYVFENASDIIEPTGDA 203  
 Db 121 DNVTKEIRNKOESHOHREGTPRNDGAVVALARSQRTYTDGYVFENADIIEDTGA 176  
 QY 204 YIVHGCHYHYIPKNELASAEALAEFLSGRNLNSRTYRRQNSDNTSRTNWPVSVN 263  
 Db 177 YIVHGCHYHYIPKSDLSASELAALAKHLAGKNNQPSQLSYSSASDN-----224  
 QY 264 PGTNTNTSNTSNTSASQSDNDISLLKQYKLPISQRYVESDGLVFPDAOTTSTARG 323

Db 225 ----NTQVAKGSTSKPANKSENLOSLLKELYDSPSAQRYSESDGLVDPAKIISRTPNG 280  
 QY 324 VAVPGDHYHFIPYQSMSELEERIAIIPLYRSHNWVDSRPEQSPQPIPEPSGPQP 383  
 Db 281 VAIPIPDHFIIPYKLSALEEKIARMV-----309  
 QY 384 APNLKIDNSSLSVQLVRRVGEVYFEKGISRYVFAKDLDESEYKLNESKLSQESVSH 443  
 Db 310 -----ISGTGVSTNAK-----PREVYSIGLSLSSNSS---339  
 QY 444 TLTAKKENAPPDQEFYKAYNLLEAKHAKLFENKGRNSDQALDKLERLNDESTNKK 503  
 Db 340 -LTTSK-----LSSA-----349  
 QY 504 LVDDLLAFLAPITPBERLQKPNQIEYTEDEYRIALQADKTYTSDGYTFEDHDIISDEGD 563  
 Db 350 -----SDGYTFNKDIYEETAT 366  
 QY 564 AVYTPHMGSHWIGKDSLSDEKVAQAQYTKKGTLPSPADAVKAP-----TGDSA 616  
 Db 367 AYIVRHGDHFYI-----PKSNQIGOPTLPNNSLATPSPSLPI--NGTSHKHEEDGY 418  
 QY 617 AAIYRNVKGEKRIPLVRLPYMVEHTVEYKNGNLLIIPKDHV 658  
 Db 419 GPDANRIAE-----DESGFVMSHGDHMH 442

## RESULT 4

US-09-134-001C-5080  
 ; Sequence 5080, Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC  
 ; FILE REFERENCE: GTC-007  
 ; CURRENT APPLICATION NUMBER: US/09/134,001C  
 ; PRIOR FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/064,964  
 ; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 5080  
 ; LENGTH: 3696  
 ; TYPE: PRP  
 ; ORGANISM: Staphylococcus epidermidis  
 ; US-09-134-001C-5080

Query Match 3.8%; Score 162; DB 4; Length 3696;  
 Best Local Similarity 18.5%; Pred. No. 0.0098;  
 Matches 183; Conservative 141; Mismatches 348; Indels 316; Gaps 42;

QY 35 NNRVSYI--DGQATQKTELTPEVSKREGINAEOIYIKITD-----QGYVTSHGDIHY 87  
 Db 1299 NNQIDIGIVSDGROSINA--ITPDTISIKRNKN--DIDIKADKKIKIQRINDATDEIQ 1353  
 QY 88 YVNGKPYDAIIEELLKMDPNYKLDKEDIVNEVGKGYIKYD-----130  
 Db 1354 EANKRIE-----EAKLEADNIDQNRSTROVNEAKTNGINKIKENTITPATYKSEARQAVQ 1408  
 QY 131 ----GKYVYLKDAADNVRKEEINR-----OKOESHOHREG--TPRNDGAVALA 177  
 Db 1409 NKANEDINIDQNPATNEKEQALNRSASAEALRAVOAQINAHHTYQGVKTIKDAITSL 1468  
 QY 178 RSGGRTYTDGTYFENASDIIEPTGDAIYVPHGDHYHYIPKNELASAEALAEFLSGGN 237  
 Db 1469 R-----INAOVEKE-----SARNA 1483  
 QY 238 LNSRTYRRQ--NSDNTS-----RTNWPVSVPNGTTNTN--TSNNSNTNSQASQSDND 286  
 Db 1484 IEOKATQQTQFNTNNDNATDEKREEVANNLVITATKQSLDNINSSLNNDVENAKVAGINE 1543





QY 777 LKNNLTLQ-----IMDNNSIMAEKLLALLKGSNPSYSKEX 814  
 Db 771 LKNNLTLQ-----IMDNNSIMAEKLLALLKGSNPSYSKEX 811

RESULT 8  
 US-09-541-782-6  
 : Sequence 6, Application US/09541782  
 : Patent No. 6284480  
 : GENERAL INFORMATION:  
 : APPLICANT: Nislow, Corey  
 : APPLICANT: Sakowicz, Roman  
 : APPLICANT: Beraud, Christophe  
 : TITLE OF INVENTION: Antifungal Assay  
 : FILE REFERENCE: 1015  
 : CURRENT APPLICATION NUMBER: US/09/541,782  
 : CURRENT FILING DATE: 2000-04-03  
 : NUMBER OF SEQ ID NOS: 10  
 : SOFTWARE: FASTSEQ for Windows Version 4.0  
 : SEQ ID NO 6  
 : LENGTH: 1073  
 : TYPE: PRT  
 : ORGANISM: S.pombe  
 : US-09-541-782-6

Query Match 3.4%; Score 145.5; DB 4; Length 1073;  
 Best Local Similarity 18.6%; Pred. No. 0.024;  
 Matches 182; Conservative 136; Mismatches 357; Indels 301; Gaps 41;

QY 43 GKATOKTENLTPPEVSKREGINAE-----QIVIKI--TDGVTSHGDHYHYNN 90  
 Db 147 GGTGCTKTYTMSGD--LSDSGILSEAGLIPRALYQLFSSLDNSNGEYAVK--CSYTELYN 204  
 QY 91 GKVPYAIISSEELMKDPNKLKDED-----IVNEVGKGYIKVDGKYVYLKDAHAHN 145  
 Db 205 EET-RDLVSEEL--RKPARVEFDTSRGNVYITGIEESY--IKNAGDGLRLREGSHRQ 260  
 QY 146 VRTKEELNRQKQHS-----QHREGTFRNDGAVALARSGRYTTDGYIFNADIIEDT 200  
 Db 261 VAATKCNLSSKSHSIFITTLHKKVSSGMDTETNSLTINN-----NSDILLR-- 307  
 QY 201 GDAYIVPHGDHYHYIPKNELASAEAFLSGRGNSLNSRT----- 243  
 Db 308 -----ASKLHWYDLAGESEN--GRSGAEKKRARETGMINOSLLTLGRVI 349  
 QY 244 -----YR-----RONSNTSRTNVPVSYPGTTNTNTSNNSTNSQASQS 284  
 Db 350 NALVEKAHHIPYRESKLTLLQDSLGKTKTSMIVTVS---STNNLEETISTLEYAARA 406  
 QY 285 NDISLKLQLYKPLSORHVESDGLVFPDAOI-----TSRTARGVAVPHGDHYHFIYISQ 339  
 Db 407 K-----STRNKQNNQVFRKYLIDYLDIERLKNDLNATRRKKGVYLAESTYKELMDRYQ 463  
 QY 340 MSEL-----EERLARIIPLRYSNHWVPDSRPEQSPQPPPEPSGQAPAPNLKI----- 389  
 Db 464 NKDLCOEARLKEVLIDLVNKS-----SREQLQYVSKSQEHEKKKEVEALQDLVNSSTE 517  
 QY 390 -----DSNSSLVSQLVAKVGEVYFPEKGSIRYFAKDLPESEYVKNLEKLSKQESVSH 444  
 Db 518 LESVSKSENEKLKELVLEIEKRRKYEFTNEAKITTVATDLSQYRESKEYIASLYEKLDR 577  
 QY 445 LPAKKENAVARDOEFYDKAVNL-----TEAH 471  
 Db 578 ERNNKEN-----ENNFWNLKFNLLTLMLRSPHSGFTDETNGFTILLDPNMSMEELLWTHNS 633  
 QY 472 KALFENKGRNSDFOALDKLLER-----LNDESTNKKRKYLDVLLAFLAPIT 516  
 Db 634 QLLISMTKITEHQSDEALQASRSCAVPNSSLDLIVSELKDSKNSLLDALEHSLQDIS 693  
 QY 517 -HBERLCKPNS-----QIETDE-----YR--IAOL 540  
 Db 694 MSSQKLTNGISSELIELQDKMKESYROLVQELSLVNLQHTHEESQKELMYGVANDIDAL 753

QY 541 ADKTYTS--DGYIPDEHDIISDEGDAYVTPHMGHSHWIGKDSLDEKYVAAQATREKGI 598  
 Db 754 VKTCTTSLNADIT--LSDYISDQSKFESKQOQDLIANIGKIVNPLQEQNESLYTK----- 808  
 QY 599 LPPSPDADYKANPTGDSAAAI--YNRVKGKRIPLVR-----LPIWEHTVEYK 645  
 Db 809 -----ADILSHLNDPTNSNIRKANEIMNNRSEEFLLNMAASQAEIYQANKERIOKTEVE-- 860  
 QY 646 NGNLIIPHKCH--YINIKFAMFDDHTYKAPNG-----YTLDELATIKYVEHDERP 696  
 Db 861 NGSQILSSKSKAHSNSRSMYDHCALAESOKQCVNLEVOITLRLQKVR-----E 911  
 QY 697 HSNQWGNASEHYLVKKDHSEDPNKNKFADEPEVEETPAPEVPOVTEKVEAQLAEAY 756  
 Db 912 HSEDNTEKKHQQLDLLESLVGNNDNL-----IDSIKTPHTELO----- 950  
 QY 757 LLAQVTDSSLK-----ANATELAGLRN-----NLTLOIMNNSIMAA-----EKL 800  
 Db 951 --KITDHVLKGTSLANHTNELGLDESLCNLETTIEDTSLVKLETTGDTPSKRELPA 1007  
 QY 801 LKGSNPSYSKEXIN 816  
 Db 1008 TPSTWROSLIKETTN 1023

RESULT 9  
 US-09-723-820-6  
 : Sequence 6, Application US/09723820  
 : Patent No. 6468760  
 : GENERAL INFORMATION:  
 : APPLICANT: Nislow, Corey  
 : APPLICANT: Sakowicz, Roman  
 : APPLICANT: Beraud, Christophe  
 : TITLE OF INVENTION: Antifungal Assay  
 : FILE REFERENCE: 1015  
 : CURRENT APPLICATION NUMBER: US/09/723,820  
 : CURRENT FILING DATE: 2000-11-28  
 : PRIOR APPLICATION NUMBER: 09/541,782  
 : PRIOR FILING DATE: 2000-04-03  
 : NUMBER OF SEQ ID NOS: 10  
 : SOFTWARE: FASTSEQ for Windows Version 4.0  
 : SEQ ID NO 6  
 : LENGTH: 1073  
 : TYPE: PRT  
 : ORGANISM: S.pombe  
 : US-09-723-820-6

Query Match 3.4%; Score 145.5; DB 4; Length 1073;  
 Best Local Similarity 18.6%; Pred. No. 0.024;  
 Matches 182; Conservative 136; Mismatches 357; Indels 301; Gaps 41;

QY 43 GKATOKTENLTPPEVSKREGINAE-----QIVIKI--TDGVTSHGDHYHYNN 90  
 Db 147 GGTGCTKTYTMSGD--LSDSGILSEAGLIPRALYQLFSSLDNSNGEYAVK--CSYTELYN 204  
 QY 91 GKVPYAIISSEELMKDPNKLKDED-----IVNEVGKGYIKVDGKYVYLKDAHAHN 145  
 Db 205 EET-RDLVSEEL--RKPARVEFDTSRGNVYITGIEESY--IKNAGDGLRLREGSHRQ 260  
 QY 146 VRTKEELNRQKQHS-----QHREGTFRNDGAVALARSGRYTTDGYIFNADIIEDT 200  
 Db 261 VAATKCNLSSKSHSIFITTLHKKVSSGMDTETNSLTINN-----NSDILLR-- 307  
 QY 201 GDAYIVPHGDHYHYIPKNELASAEAFLSGRGNSLNSRT----- 243  
 Db 308 -----ASKLHWYDLAGESEN--GRSGAEKKRARETGMINOSLLTLGRVI 349  
 QY 244 -----YR-----RONSNTSRTNVPVSYPGTTNTNTSNNSTNSQASQS 284  
 Db 350 NALVEKAHHIPYRESKLTLLQDSLGKTKTSMIVTVS---STNNLEETISTLEYAARA 406  
 QY 285 NDISLKLQLYKPLSORHVESDGLVFPDAOI-----TSRTARGVAVPHGDHYHFIYISQ 339

```

Db 407 K---SIRNKKQNNQOLVFRKLLINDVLDIERLNDLNATRRKNGVYLAESTYKELMDRVQ 463
QY 340 MSEL---EERIRARILPLRYRSHHWVDSRPEQSPQPTPEPSPQAPANLKI----- 369
Db 464 NKDLLOCOARKLEVDLANKS-----SFEQJOYVSKSQEHEKKEVEALQJOLVNSSTE 517
QY 390 ----DSNSSLVSQLYKRVGEGYVFEKGISRYVFAKDLPSETVKMLSKESKQESVSHT 444
Db 518 LESVKSSENEKLNELVLEIEKRKKYEENAKITVTATDLSQYREKRYIASIYEKLDRT 577
QY 445 LTKKEVAVARDEPFYDKAVNL-----TEAH 471
Db 578 ERNNKEN---ENNFWNLFFNLLTMLRSPHGSFTDETNGYFTLLDNPNASMEELLNTHSN 633
QY 472 KALFENKGRNSDFQALDKLER-----LNDESNKREKIVDDLLAFLAPT 516
Db 634 QLLISMTKITEHQSIDENALQSRSSCAVPNSSLDIYSELKDSKNSLIDLALHSHLQDTS 693
QY 517 -HPERLGKPN-----QIEYTEDE-----VR--IAOL 540
Db 694 MSSQKLGNGISSELILOKDMKESYQOLVQELRLYLNQHTHESEQELMYGRNDIDL 753
QY 541 ADKYTS---DGYTFEDHDIISDEGDAYVTPHMGSHWIKRDSLSKEKYAAQATYKSGI 598
Db 754 VKCTISLMDADII-LSDYISDQKSKFESKQODLIANIKRIVSNFLQEQNESILYTK--- 808
QY 599 LPPSPADYKANFTGDSAAI--YNRVKGKRIPLVR-----LPVWEHTVEYK 645
Db 809 ----ADLHSHLNDTNSNIRKANEMNNRSEELRNAAOAEIVGANKERLOKVE--- 860
QY 646 NGNLIIPKDH-YHNKIFAMFDQHTYKAPNG-----YTLEDLFAITKIYVHDEPER 696
Db 861 NGSQLDLSKSKAIHSRSRMTDHCIALAESOKGVNLEVOTLDRLLQYK-----E 911
QY 697 HSDNGMNASEHVLCRKHSEDPNKNFKADEBPVEETPAPEPVQVETKEVFAQLKEAV 756
Db 912 HSEDNTRKHHQOLDLDESIVGNNDL-----IDSKITHTLQ----- 950
QY 757 LAKVYDSSLK-----ANATETLAGLRN---NLTLQIMDNNSIAMA-----EKLLA 800
Db 951 ---KITDHALKGTSLANTNELLGLGDSLCNLETTJEDTSLVLETTGTDPSPREKLPA 1007
QY 801 LKGSNPSVSSEKIN 816
Db 1008 TPSWTRDSSLIKETTN 1023

RESULT 10
US-08-961-083-118
; Sequence 118, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: ChOI et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 118:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1040 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-961-083-118

```

Query Match 3.4%; Score 144; DB 4; Length 1040;

Best Local Similarity 19.2%; Pred. No. 0.03; Mismatches 284; Indels 252; Gaps 38;

```

Matches 154; Conservative 111; Mismatches 284; Indels 252; Gaps 38;

QY 127 IKVDGKYV-YLKDAHADNVRKEELNRQKQESQHREGGTPRND-----G 172
Db 76 LKIEGYQYIGYIK-TKQDNTELSTRTVDGKYSQGRDSQPNSTKTSVDYVHSDLEMGOG 134
QY 173 AVALARSGRYTTDGGYIFNASDIIEDTGDAYIYPHGDHYIIRKNELSASELAALAPL 232
Db 135 KVSU---QGEASGDDG-----LSKSSIAADNLSSNDSFA 166
QY 233 SGRNLMSRTYRROSNDTSRTMWPVSNSPGTNTNTSNTNSNTNSOASNDIDSLLK 292
Db 167 S-----QVEQNPDKHGESVVRPLVPEQGNPVSATVQSAEEVYLATND---R 211
QY 293 QLYKPLSQRHVESDG-----LVF-----DPAQITS 318
Db 212 PEKLPLELTQTQEBGHEGAARVDELPVYKPLETKGTGQPGHGEAAVAREEPATYEP 271
QY 319 RTARGVAVPHGDHYHTFYPSQMSLEERARILPLRYRSHHWVDSRPEQSPQPTPEPS 378
Db 272 LATGTQEPGHE-----GKATVREELLETTERPVATKGT---QEPHEGEXVVEEL 319
QY 379 PGQPA-----PNL-----KIDNSSSLVQLVR--KVGEGYFEE 411
Db 320 PALEVTYRNRTEIONIPYTEELIDPILLKNRKLEKQAGCTRTQYEDYIVANGVET 379
QY 412 KGISRYFAKDLPSETVKYNLESKLSKQESVSHLTAKKENAVARPDQFYDKAVNL--TE 469
Db 380 KEVSRTEVA---PVNEVYKGTIVKVPYVETINLTKVEN---KRSIVSYNLLIDTTS 431
QY 470 AHKALFENKGRNSDFQALDKLERLNDSTNKEKLYVDLLAFLAPI---TH-PERLGKPN 525
Db 432 AYVS-----AKTQVFHG-DKLVKEVDIENPAKEVIGS-LDYITPYVYKHLHYLNQENN 484
QY 526 SQIYTEDEVRIADLAKYITTSQGITDEHDIIISDEGDAYVTPHMGSHWIKGDSLSKE 565
Db 485 E--INTETSTQDFOL--EYKKEIKIDISVELYKENDRY-----RRYL---SISEAP 530
QY 586 KVAQAYTKEKGLPSPDADYKANPTGDSAAIYNNVKGKRIPLVRILPYWEHT--- 631
Db 531 TDTAKYFVYK-----SDRFKEVYLPVKS-----ITENTDGYT 563
QY 642 -VEYKNGNLIIPKDHYNIRKFAVFDHYT-----KAPNGYT-LEDLFAITK-----Y 687
Db 564 KVVYAVADOLVEEGTDGK-----DDYTFYVAKSKAEQGVGYTSFKQLVTAMQNSLGV 616
QY 688 YVEHPDERPHSNQGWGASHEVLCRKHDS-----EPNKNKRADE--EPVEET 733
Db 617 YTLASDM-----TADVSLGDQTSYLTGAFGSLIGSGTGSYAIYDLKPLPDT 667
QY 734 PABEVPQVETKEVADQKEAEVLLAKVTDSL-----KANATEETLAGLNNNTLQI 785
Db 668 LNGATVARDLDIKTVSADSKENVAALAKKANSANINNVAVGKISGAKSVAGLVASATNTV 727

```

QY 786 MDNNSIMAEKLLALLKGSN 806  
:1:1 :1:1 :1:1  
Db 728 IENSFTG---KLINHQDSN 745

RESULT 11  
US-09-816-703A-2  
Sequence 2, Application US/09816703A  
Patent No. 6455026  
GENERAL INFORMATION:  
APPLICANT: Melcher, Thorsten  
APPLICANT: Mueller, Sabine  
APPLICANT: Chin, Daniel  
TITLE OF INVENTION: Use of Protein Tyrosine Phosphatase Zeta as a Biomolecular Target  
FILE REFERENCE: 262/235 AGY  
CURRENT APPLICATION NUMBER: US/09/816,703A  
CURRENT FILING DATE: 2001-03-23  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 2314  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: Reference  
LOCATION: (1)..(2314)  
OTHER INFORMATION: Krueger, N.X. and Saito, H.: A human transmembrane protein-tyrosi  
OTHER INFORMATION: ne-phosphatase, ptp zeta, is expressed in brain and has an N-term  
OTHER INFORMATION: 1n1 receptor domain homologous to carbonic anhydrases; Proc. Nat  
OTHER INFORMATION: L. Acad. Sci. USA 89 (16), 7417-7421 (1992)  
NAME/KEY: Reference  
LOCATION: (1)..(2314)  
OTHER INFORMATION: Levy, J.B., et al.: The cloning of a receptor-type protein tyrosi  
OTHER INFORMATION: ne phosphatase expressed in the central nervous system; J. Biol.  
OTHER INFORMATION: Chem. 268 (14), 10573-10581 (1993)  
NAME/KEY: SIGNAL  
LOCATION: (1)..(24)  
OTHER INFORMATION: By similarity  
NAME/KEY: DOMAIN  
LOCATION: (25)..(1635)  
OTHER INFORMATION: Extracellular (potential)  
NAME/KEY: misc\_feature  
LOCATION: (25)..(2314)  
OTHER INFORMATION: Mature chain; protein-tyrosine phosphatase zeta  
NAME/KEY: DOMAIN  
LOCATION: (34)..(302)  
OTHER INFORMATION: Carbonic-anhydrase like  
NAME/KEY: SITE  
LOCATION: (105)..(105)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: SITE  
LOCATION: (134)..(134)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: SITE  
LOCATION: (223)..(223)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: SITE  
LOCATION: (232)..(232)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: DOMAIN  
LOCATION: (312)..(406)  
OTHER INFORMATION: Fibronectin Type-III  
NAME/KEY: SITE  
LOCATION: (324)..(324)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: SITE  
LOCATION: (381)..(381)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: SITE  
LOCATION: (497)..(497)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)

NAME/KEY: SITE  
LOCATION: (501)..(501)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: SITE  
LOCATION: (552)..(552)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: BINDING  
LOCATION: (587)..(587)  
OTHER INFORMATION: Chondroitin sulfate (potential)  
NAME/KEY: SITE  
LOCATION: (602)..(602)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: SITE  
LOCATION: (629)..(629)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: BINDING  
LOCATION: (637)..(637)  
OTHER INFORMATION: Chondroitin sulfate (potential)  
NAME/KEY: SITE  
LOCATION: (677)..(677)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: VARSPLIC  
LOCATION: (755)..(1614)  
OTHER INFORMATION: Splicing variant; missing (in short isoform)  
NAME/KEY: BINDING  
LOCATION: (997)..(997)  
OTHER INFORMATION: Chondroitin sulfate (potential)  
NAME/KEY: SITE  
LOCATION: (1017)..(1017)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: SITE  
LOCATION: (1050)..(1050)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: SITE  
LOCATION: (1082)..(1082)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: SITE  
LOCATION: (1122)..(1122)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: SITE  
LOCATION: (1456)..(1456)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: BINDING  
LOCATION: (1548)..(1548)  
OTHER INFORMATION: Chondroitin sulfate (potential)  
NAME/KEY: BINDING  
LOCATION: (1550)..(1550)  
OTHER INFORMATION: Chondroitin sulfate (potential)  
NAME/KEY: SITE  
LOCATION: (1561)..(1561)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: SITE  
LOCATION: (1617)..(1617)  
OTHER INFORMATION: Glycosylation; N-linked (GLCNAC...) (potential)  
NAME/KEY: TRANSMEM  
LOCATION: (1636)..(1661)  
OTHER INFORMATION: Transmembrane region; potential  
NAME/KEY: DOMAIN  
LOCATION: (1662)..(2314)  
OTHER INFORMATION: Cytoplasmic (potential)  
NAME/KEY: CONFLICT  
LOCATION: (1722)..(1728)  
OTHER INFORMATION: Missing (in ref. 2)  
NAME/KEY: DOMAIN  
LOCATION: (1744)..(1997)  
OTHER INFORMATION: Protein-tyrosine phosphatase  
NAME/KEY: ACT\_SITE  
LOCATION: (1932)..(1932)  
OTHER INFORMATION: Active site; by similarity  
NAME/KEY: DOMAIN  
LOCATION: (1998)..(2314)  
OTHER INFORMATION: Protein-tyrosine phosphatase  
NAME/KEY: misc\_feature

LOCATION: (2222)..(2222)  
OTHER INFORMATION: Ancestral active site  
US-09-816-703A-2

Query Match 3.3%: Score 142.5; DB 4; Length 2314;  
Best Local Similarity 19.2%; Pred. NO. 0.14; Mismatches 267; Indels 331; Gaps 43;  
Matches 169; Conservative 113; Mismatches 267; Indels 331; Gaps 43;

```

QY 71 IKITDOGVYVTHSHGHYHYNGKVPYDAIISPELLMKDPNKKLKDEDIVNEVKGVIKVD 130
DB 954 VGVTVQSGSLFSGPSHI-----PIPKSLITPTASLQPTHLSG-----D 993
QY 131 GKYYVYLKDAHADNVKKEIKINROKQEHSHRGGRGTFRNDGAVL-----ARSQRYT 184
DB 994 GEM-----SGASSDS-----EFLLPDTGDLALNITSPSVAAEFTYT 1030
QY 185 T-----DQGYTFNADSLIEDTDAIVPHGDHYH-----IPK-----NELSASELAAE 229
DB 1031 TSVEGDDKALKSKSEITIGNETELQIPSEFNMYPSESTVPRNMYDNVKNLNASLOETSV 1090
QY 230 AFLSGRG-----NLNSRTYR--RONSNDNTS-----RTNWVP 259
DB 1091 SISRKCMFPGSLAHTTKVFDHEISQVPENNFSQVPTHTVSQASGDTSLKPVLSANSEP 1150
QY 260 SVSNPGR-----NTNNSNNTNSQAS--QSDIDSLKQLYKPLPSORHVES 306
DB 1151 ASSDPASSEMLSPSTOLFEYTSASFSTEVLLQPSFQASDVDTLKVY--LPA-----VPS 1204
QY 307 DGLVDPDPAQ-----ITSRTARGAVPHGDHYHFIYQMSSELEERATIPRY 355
DB 1205 DPLIVETPKVDKISSTMHLIVSNSASENNLHSTSVPEFVSPTSHHNSASLOGLTISTY 1264
QY 356 RSNHWVPDSREQSPQPTPEPSPGQPPAPNLKIDSNSLSVQLVKKGSEYVEEKGIS 415
DB 1265 ASEKYEPLKSESQHYVP-----SLYSNDELQTNLEINQAH--PKG-- 1308
QY 416 RYVFAKDLPS--EYVKNEESKLSKQESVSHLLAKKENVADR-----DQEPYDKAY 464
DB 1309 RHVFATPVLSIDEPNLINKLINSDEL--LTSKSSVTKGVFAGIPTVASDFVSTDH 1365
QY 465 NL-LFEAH-----KALFENKG----- 479
DB 1366 SVPIGNHVAITAVSPHDSVSTKLLFPSKATSELSHSAKSDAGLVGGEDGDDTDDG 1425
QY 480 -----RNSDFOALDKLERLNDSTENKELVDDLFLAPLTHPERLGPMSQIEY----- 530
DB 1426 DDDDDROSDGLSIHKCMS--CSSYRESQEKVMD-----SPTHEMSLMDQNNPISYSLSE 1478
QY 531 -TEDEVRIQLADKYTT-----SDGYTFDEHDI-----ISDEG 562
DB 1479 NSEEDNRYTSVSSDSQOTGMDRSPGKSPSANGLSQKHNDGK--EENDIQTGSAALLPLSPES 1536
QY 563 DAYVVRPHNGHSHWIGK--DSLQKEKVA--AQAYTKRK--GIL-----PPSP 603
DB 1537 KAMAVULIGDESQSGQSTSDSLNENETSTDFSPADTNKEDADGILLAGDSITPGFPQSP 1566
QY 604 DADV-----KANPTGSAAIYNRVYKGEKR--IPLVRLPYMEHTVEVANG 647
DB 1597 TSSVTSENSEVFVHSEAFASNSHESRIGLAEGLSEKKAIVPLIYVSALFFLGLVLYG 1656
QY 648 NL-----IIPKHDIYHNKFAWEDDH--YKAPNGY 676
DB 1657 ILIYWRKCFOTAHFLEDSTSPRVISTPPIFIPISDVGAIPIKHFPKHVADLHASGCF 1716
QY 677 TLEDLFATIK--YVVE-----HPPERPHSN 699
DB 1717 TEE--FETLKEFYQVQSVYDGLITADSSNHPDK--HKN 1753

```

RESULT 12  
US-08-976-255-14  
Sequence 14, Application US/08976255  
Patent No. 6136581

GENERAL INFORMATION:  
APPLICANT: Jono, Keith E.  
APPLICANT: Plowman, Gregory  
TITLE OF INVENTION: KINASE GENES AND USES  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: Suite 4700  
City: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071-2066  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: FASTSEQ for Windows 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/976,255  
FILING DATE: NO. 6136581ember 21, 1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/031,675  
FILING DATE: NO. 6136581ember 22, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Warburg, Richard J.  
REGISTRATION NUMBER: 32,327  
REFERENCE/DOCKET NUMBER: 229/182  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ. ID NO.: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1503 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-976-255-14

Query Match 3.3%: Score 140; DB 4; Length 1503;  
Best Local Similarity 19.4%; Pred. NO. 0.11;  
Matches 150; Conservative 119; Mismatches 275; Indels 228; Gaps 38;

```

QY 54 TPDEVSKREGINAEQIYKTTDOGVYTHSGD-----HYHYNG-KVPYDAIISELL 104
DB 582 TGPESLQITAIRSEVLEESSTDEDFFQSSTPDKOSSLPGLDHTSGPESFFNIF----- 636
QY 105 MKDPNRYKLKDEDIVNEKGGYVIRKVDGKYVYLKDAHADNVKKEIKINROKQESHORE 164
DB 637 -----NDVKSIEDLPISHQKTFDLMLNG--VOADTKPATLSSLD-----NPKESVITGHE 686
QY 165 GGTFRNDGAVALARSGQRYTTDGYIFNASDI--TEDTGAVIVPHGDHYHYPKNELSAS 223
DB 687 KEKRRK-----IFDSEPLCLSDN-----LMHQDNF-----DPLNVO 717
QY 224 ELAAAEFLSGRGLNSKRYRNRONSNDNTSRTWVSVSNPCTT-----NTNNSNNTNS 279
DB 718 ELSENFLFQEKMLKGLSSKKEHINDLP-----ELKNAGFTEMLTSCNSISDTL 771
QY 280 QASQNDIDSLKQLYKPLPSORHVESDG-----LVFPDPAQITSRARGVAVPHGDHYH 333
DB 772 QFAENKRGSL-----QENVSTKGDDTVMLTGLTSLSSSPVQVY----- 816
QY 334 FIYQMSSELEERIALIPLRYRSNHWVPDSREQSPQPT-----PEP-----SPGP 381
DB 817 --PTS--FETEETPRRY-----PPDSLPTQGTQCTCDVIVPEDCQLHQDISPA 862
QY 382 QPAPNLKIDSNSLSVQLVKKGSEYVEEKGISRYVFAKDLFSEYVKNESLSKQESV 441

```

```

Db 863 VTPV-VEILSTDARTSLDNR-----SODSPGESEETL--RLTESDSV 902
QY 442 -SHTLTAKKENVAPR-----DOEFYKAVNLJTEAHKALPENKGRMSDQALDK----- 489
Db 903 LADDILASVSVGSSLPGLGOLHNKPFSEDSHSHRLEKLNLEAVETLNOLNSKAAEA 962
QY 490 -LLERLNDESTKEKLVDDLLAFAPITHPERLGPNSQIEYTEDEVRILAOLADKYTSD 548
Db 963 GLVSLSDSTQSODSLLEDLSLAPPPASEPS-LETTPDS-LESV----- 1003
QY 549 GYIFPEHDIISDEGDAYVTPHMGSHWIGKDSLKKEVAAQATYKE-----KGLTPSP 603
Db 1004 ----DVHEALDLSLGH-TP-----OKLVPPDKPAOSGYETENLESPEWTLHPAP 1048
QY 604 DADVAKNP--TGDSAAATYNNRYKGEKRIPLVRLPYMVEHTVEVKNGLIIPKHGHNIK 661
Db 1049 EETADEPRTTGDGSHS-----GLPPNPVI-----VIDAGGHGTEVTP----- 1089
QY 662 FAMPDDHYKAPNGYTLLEDLFATITYVVEHPDERPHSNDGMCMAEHVGLKKHSEDPNK 721
Db 1090 -----ETFTAGSQCSYRD-----SAVFSNDSEPEKR-----SEEVPG-----TSPSA 1127
QY 722 NKADEEPVEETPAPEVYQVETEKEVQAOLKEAF--VLLAKYTDLSLKANAT 771
Db 1128 LVLVQEQPLPEVLPPEQSPAQDSCLAEARKSQPDESCLALHNSDLELRAT 1179

```

# RESULT 13

US-07-920-281C-2

```

; Sequence 2, Application US/07920281C
; Patent No. 5739026
; GENERAL INFORMATION:
; APPLICANT: Garoff, Henrik
; APPLICANT: Liljestrom, Peter
; TITLE OF INVENTION: DNA Expression Systems Based on
; TITLE OF INVENTION: Alphaviruses
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/920.281C
; FILING DATE: 13-AUG-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 828-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-920-281C-2

```

Query Match 3.3%; Score 140; DB 1; length 2431;  
 Best Local Similarity 19.7%; Pred. No. 0.23;  
 Matches 137; Conservative 93; Mismatches 253; Indels 212; Gaps 34;

```

QY 76 QGYVTSBGDHYHYNG-KVEPYDAIISELLMKDPNKLKDEEDVINEKGGYIKVDGKYY 134
Db 1523 KGYETTTGSLSYTEGKRFQNAALDMAEILLTLPRLQEANBQICVIALGTM----- 1574
QY 135 VYLDAAHADNVKTKKEIINQKOBHSQHREGGTPRNDGAVALARSQGRYTTDDGYIFNAS 194
Db 1575 -----DNIRSKCYVN-----DSDSSTPPRT-VPCLCR-----YAMTAE 1606
QY 195 DIEDTG-----DAYIVPHGDHYHYIPKNELASLELAAEAFISGNGLSNRTAR 245
Db 1607 RIALLRSHQKSNWVCSSFLP--KYHVGQVKCEKVLLEDPVY--SVYSPKKA 1660
QY 246 RQNSDNTSRT-----NWVSVSNPGTTNTNTSNSTNSQASQSDIDSLKQLYKLP 299
Db 1661 ASTTDHSDRLRGFDLW-----TTDSSSTASDTMSLPISQSDIDISYEPMAPIV 1712
QY 300 -SQNHVSDGLVEDPQAOTISRTAGVAVPBGDHYHFLYPSQSELEERLARIILPLRSN 358
Db 1713 TADVHPAPAGIADLAADVHEPA-----DHV-----DLENPIV--PPRPKRA 1752
QY 359 HWPVDSRPQPPQPPPEPSPGPPAPNLKI-----DSNSLSVQSLVRKVGEGYVEEK 413
Db 1753 AYLSRAAEKRPVAPR-KPPAPRTAFRANKPLTFGDFDEHEVDALASGITTSDEDDVLR 1811
QY 414 ISR---YVFAKDLPESETVKNLESKLSKQESVSHLLAKKENVAPRQDEYDKAYNLTEA 470
Db 1812 LGRAGAVITFSSDGS---GHLOQKSVRQ---HNLOCAQUDANAQOEKMPPK---LDTER 1861
QY 471 HKALF-----ENGRNSDFQALDKLLERLNDESTKEKLVDDLLAFAP---ITHP 518
Db 1862 EKLLLMQMHPSSEANKSRQS-RKVENMKATYVDRLTSGARLYTGADVGRIPTYAVRYP 1920
QY 519 ERLGPNQSLQTEYDEVRIA---QLADKYTSDGY-IFPEHDIISDEGDAYVTPMGHS 573
Db 1921 RPYVSPVIERFSSPDVAIACNEYLSRNPPTYASQITDEYATYDMD----- 1970
QY 574 HWIGKDSLDE---KVAQAATYKGT-----LPPSP-----DADYKAN----- 610
Db 1971 ---GSDSCLDRATFCPAKLCYRHNHAYHQPVTYSAVPSFQNTLQNLVLAALAKRKCNT 2027
QY 611 -----PTGDSAAATYN-----RVKGEK----- 628
Db 2028 QMRLEPTMDS--AVFVNECKRACSGEYWEYAKODIRITENTITTYTKLGPAAAL 2085
QY 629 -----IPLVRLPYMVEHTVEVKNGLIIPKH 656
Db 2086 FAKTHNLVPLQEVN-MDRFTVDMKRDYKVTPTGKH 2119

```

# RESULT 14

US-08-466-277-2

```

; Sequence 2, Application US/08466277
; Patent No. 6190666
; GENERAL INFORMATION:
; APPLICANT: Garoff, Henrik
; APPLICANT: Liljestrom, Peter
; TITLE OF INVENTION: DNA Expression Systems Based on
; TITLE OF INVENTION: Alphaviruses
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,277

```



```

; FILING DATE: 06-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/920,281
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 828-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-466-277-2

```

```

Query Match          3.3%; Score 140; DB 4; Length 2431;
Best Local Similarity 19.7%; Pred. No. 0.23; Indels 212; Gaps 34;
Matches 137; Conservative 93; Mismatches 253;

```

```

QY 76 QGYTSHGDHYHYNG-KVYPDALISEELMKDPNYKLKDEIVNEVKGQYVIRKDKY 134
DB 1533 KGYSTTDSLSYFEGTKFNQADIMAEILTMPRLQANQICLYALGETM----- 1574
QY 135 VYLKDAHADNVRTKEELNROKQSHOREGCTPRNDGAVALANSQGRYTTDDGYITNAS 194
DB 1575 -----DNRKSCPVN-----DSDSTPRT-VPCLCR-----YAMTAE 1606
QY 195 DIEDTG-----DAYIVPHGDHYHYIPKNELSASELSAAEAFISGRGNLSNSTYR 245
DB 1607 RIARLSHQVKSVMYVCSFLP---KYHVDGVQVKCEKXVLEPPIY---SVVSPKIA 1660
QY 246 RQNSDNTSRT-----NWVSVSNPCTTNTNTSNTNSOASOSNDIDSILKQLYKPL 299
DB 1661 ASTDHSRSLRGFDLW-----TTDSSSTASDTMSLPSLSQSCIDSIYEPMAPIV 1712
QY 300 -SQRHVESDGLVFDPQAITRTARGAVAPPHGDHYHFIPIYSQMSLEIRIARIIPLRSN 358
DB 1713 TADVHPBPAGIADLAADVHPEPA-----DHV-----DLENPIIP---PPRKRA 1752
QY 359 HWVDSRPEQSPQPTPEPSPGQAPADNLKI-----DSNSLSVSLVRYKVGEGYVEEKG 413
DB 1753 AYLASRAAERBPVAPR-KPTPAPRTAFRNKLPLEFGDFDEHEVDALASGITFGDPLDVL 1811
QY 414 ISR---VVFAKDLPESEVTKULESKLSKQESVSHTLPAKKENVAPROEFYDKAYNLLEA 470
DB 1812 LGRGATITFSSDGS---GHLQOKSVRQ---HNLQCAQIDAQVQEKMPPK---LDTER 1861
QY 471 HKALF-----ENKGRNSDFQALDKLERLNDESTNKEKLVLDLAFAPL---ITHP 518
DB 1862 EKLILLKMQHPSEANKSRYS-QS-RKVENMKATVYVDLISGARKLYTADGRIPTYAVRIP 1920
QY 519 ERIGKPRQSQLETEDEVRIA-----QLADKTYTSDGY-IEDEHDIISDEGDAYVTPMGHS 573
DB 1921 RPYVSPVYIERFSSPDVAIAACNEYLNRNPVYASQYIDEDAYDAMDV----- 1970
QY 574 HWIGKDSLSDKE---KVAQAQYTKKEGI-----LPSPS-----DAVYKAN----- 610
DB 1971 ---GSDSCDRATFCPAKILRCYKHNHAYHQPYRSAPVSPFQNTLONVLAATKRKNVNT 2027
QY 611 -----PTGSAALIV-----RYKGEKR----- 628
DB 2028 QMBELPTMDS-AVFVNECFKRYACSGEWEYAKQPIRITTENTITTYTATKLGPRAAAL 2085
QY 629 -----IPLVRLPYVVEHTVEVKNGLIIPKDH 656
DB 2086 FAKTHNLVLPQEVF-MDRFTVDDKKRDKVAVTPGSKH 2119

```

```

RESULT 15
US-09-513-783a-22
; Sequence 22, Application US/09513783A
; Patent No. 6416959
; GENERAL INFORMATION:
; APPLICANT: Giuliano, Kenneth A.
; APPLICANT: Kapur, Ravi
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-L1
; CURRENT APPLICATION NUMBER: US/09/513,783A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ. ID NOS.: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1610
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: EYFP-DEV-D-MAP4-EBFP construct
US-09-513-783a-22

```

```

Query Match          3.2%; Score 137.5; DB 4; Length 1610;
Best Local Similarity 19.8%; Pred. No. 0.19; Indels 265; Gaps 41;
Matches 174; Conservative 103; Mismatches 337;

```

```

QY 73 ITDQGY-----VTSHGDIHYHYNGKVPYDAILISEELMKDP-NYKLKDE-----DIV 118
DB 62 VTTFGGYGLQCFARYPDMHQHDFKSAFMEGYQERTIFFKQDGNKTAEVKFEQDITLV 121
QY 119 NEVY-KGYVYKVDG-----KYYVYLKDAHADNVRTKEELNROKQSHS---Q 161
DB 122 NRLEKLGIDPKEDGNILGKLEVNYSNHWYIMADKQNGIKVFNKIRHNIEDSGVQLAD 181
QY 162 HREGCTPRNDGAV-----ALARSQG-----RYTTDDCY-----IFANS 194
DB 182 HYOQNTPIDGQPLRPDLNHYLSQVSLSKDPNEKRDHWLFEVYIAGITTLGMDLYKKG 241
QY 195 DIEDTGDAIYIPHGDIHYIP---KNELSASELSAAEA-----FLS 233
DB 242 DEVDGMADSLV---DALTEPPEIEGELIKRDMALAEAPYDIDVGETVEKTEFPIPLD 298
QY 234 GRGNLSRTRYRQNSDNTSRNHWVSVNPCTTNTNTSNTNSOASOSNDIDSILKQ 293
DB 299 GDEKTNSSSKKPKCLD-TSQVEGIPS-SKP-TLLANGDHGEGNNTAGSPTD---FLEE 352
QY 294 LYKLP---LSQRHVESDGLVFDPQAITRTARGAVAPPHGDHYHFIPIYSQMSLEIRIARI 350
DB 353 RVDYEDYQSSQWMPEDASFCFQPGQVLDTDQ---AEPFNEH-----RDDGLADL 398
QY 351 IPLRYSNHWVDSRPEQSPQPTPEPSPGQAPADNLKIDSNSLSVSLVRYKVGEGYVE 410
DB 399 LFVSSGPTNASFTEKIDNS-----EDSYGMLPQCSF---ASTAVSOEWSVGAPNSPCS 450
QY 411 EKGISRYFAKDLPESEVTKULE---SKLSKQESVSHTLPAKKENVAPROEFYDYDAYN 465
DB 451 ESCVS-----PVTIETIDPATELSKAAEVESVKQDLAK----- 485
QY 466 LITTEHAKLFENKGRNSPQALDKLERLND-----ESTNKEKLVLDLAFAPITHPERL 521
DB 486 -----ALETMAEQTTDVVHSPSDTTPPGPDTEALAKDLE----- 520
QY 522 GKPNQIETVEDEVRIADLADKYTTSQGYIDDEHDIISDEBDAY----- 565
DB 521 -----ETKPDVILANVTOPSTESDMFTLAQDDELLGTGAALANNIILPTEPDESSTK 573
QY 566 -VTPMGHSHWIGKDSLSDKEVAA-----QAYTKREGILPPSPDADVANKAPT 612
DB 574 DVAPMEEEIIVGNDTSPKTEFTTLPRIKMDLAPDEVVLTKETELAP---AKGNVSL 629
QY 613 GDSAAAIYNNRYKGEKRIPLVRLPYVVEHTVEVKNGLIIPKDHYNIKFAMWDDHTYKA 672

```

```

Db      630 EIEPALAKNDVRSAE-IPVAQ-ETVVSETEVVLATEVLP----- 667
QY      673 PNGYTLDELFAIKYVEHP--DERPHSNDQWGN-ASEHVLGKKD-HSEDPNKNFKADEE 728
Db      668 -----SDPTTLTKDVTLLPLEAERPLVTDMTPLSETEMTLGKETAPPTETNIGMAKDS 721
QY      729 PVEETPAE-----PEVPQVETEKVEAOLKEAEVLLAKVTDSLSKANATEVLGIRNN 780
Db      722 PLPESEVTLGKDVVILPETKVAEFNNV-TPLSEEEV--TSVKDMSPSA---ETEAPLAKN 775
QY      781 LTLO-----IMDNNSIMAEAEKLLALLKGSNPSSVSKEK 814
Db      776 ADLHSGTELLIVDNS--MAPASDLALPLETKVAIVPIKOK 812

```

Search completed: May 19, 2003, 08:25:07  
 Job.time : 32.7448 secs

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 Compen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 07:51:03 ; Search time 234.436 Seconds  
(without alignments)  
335.715 Million cell updates/sec

Title: US-09-645-835a-7

Perfect score: 4265  
Sequence: 1 MKINKRYLVGSAALILISVC.....KLALLKGNSVSKKIN 816

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 362588 segs, 96450795 residues

Total number of hits satisfying chosen parameters: 362588

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubppaa/PC1\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubppaa/PC1US\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4163	97.6	796	10	US-09-765-272-56
2	3304	77.5	840	9	US-09-884-465A-7
3	2832.5	66.4	838	9	US-09-884-465A-8
4	2661.5	62.4	763	10	US-09-765-272-66
5	1902	44.6	1126	9	US-09-884-465A-383
6	1894.5	44.4	1365	9	US-09-884-465A-382
7	1798.5	42.2	1139	9	US-09-884-465A-380
8	1798.5	42.2	1378	9	US-09-884-465A-378
9	1668	39.1	1238	9	US-09-884-465A-381
10	1659	38.9	999	9	US-09-884-465A-376
11	1652	38.7	999	9	US-09-884-465A-377
12	1295.5	30.4	1039	9	US-09-884-465A-6
13	1295.5	27.8	447	10	US-09-765-272-182
14	981.5	23.2	913	9	US-09-884-465A-384
15	991.5	23.2	1152	9	US-09-884-465A-379
16	975.5	22.9	840	9	US-09-884-465A-10
17	927	21.7	793	9	US-09-252-088-15
18	724.5	17.0	906	9	US-09-884-465A-369
19	715.5	16.8	906	9	US-09-884-465A-373

20	711.5	16.7	906	9	US-09-884-465A-371	Sequence 371, App
21	676.5	15.9	900	9	US-09-884-465A-333	Sequence 333, App
22	660	15.5	715	9	US-09-252-088-16	Sequence 16, App1
23	645	15.1	906	9	US-09-884-465A-370	Sequence 370, App
24	645	15.1	906	9	US-09-884-465A-372	Sequence 372, App
25	643	15.1	272	9	US-09-884-465A-258	Sequence 258, App
26	643	15.1	895	9	US-09-884-465A-344	Sequence 344, App
27	643	15.1	901	9	US-09-884-465A-343	Sequence 343, App
28	641	15.0	272	9	US-09-884-465A-300	Sequence 300, App
29	641	15.0	272	9	US-09-884-465A-301	Sequence 301, App
30	636	14.9	894	9	US-09-884-465A-336	Sequence 336, App
31	636	14.9	900	9	US-09-884-465A-335	Sequence 335, App
32	636	14.9	906	9	US-09-884-465A-332	Sequence 332, App
33	634	14.9	272	9	US-09-884-465A-293	Sequence 293, App
34	634	14.9	272	9	US-09-884-465A-294	Sequence 294, App
35	634	14.9	272	9	US-09-884-465A-295	Sequence 295, App
36	627	14.7	894	9	US-09-884-465A-340	Sequence 340, App
37	627	14.7	894	9	US-09-884-465A-342	Sequence 342, App
38	627	14.7	900	9	US-09-884-465A-339	Sequence 339, App
39	627	14.7	900	9	US-09-884-465A-341	Sequence 341, App
40	626	14.7	270	9	US-09-884-465A-302	Sequence 302, App
41	625	14.7	272	9	US-09-884-465A-296	Sequence 296, App
42	625	14.7	272	9	US-09-884-465A-298	Sequence 298, App
43	625	14.7	272	9	US-09-884-465A-346	Sequence 346, App
44	625	14.7	895	9	US-09-884-465A-345	Sequence 345, App
45	625	14.7	901	9	US-09-884-465A-345	Sequence 345, App

#### ALIGNMENTS

RESULT 1  
US-09-765-272-56  
Sequence 56, Application US/09765272  
Patent No. US2002006155A1  
GENERAL INFORMATION:  
APPLICANT: Choi et. al.  
TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines  
NUMBER OF SEQUENCES: 452  
CORRESPONDENCE ADDRESS:  
ADDRESS: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,272  
FILING DATE: 22-Jan-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/961,083  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, A. Anders  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: PB340P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 796 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 56:

US-09-765-272-56

Query Match 97.6%; Score 4163; DB 10; Length 796;  
 Best Local Similarity 99.9%; Pred. No. 8 2e-255;  
 Matches 795; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 21 SYELGLYQARTVKEENNVSVIDKQATQKTENLTPEVSKREGINAEQIYIKITDQGYVT 80  
 |||||  
 Db 1 SYELGLYQARTVKEENNVSVIDKQATQKTENLTPEVSKREGINAEQIYIKITDQGYVT 60  
 QY 81 SHGDHYHYNGKPYDAIIESELLMKDPNKLKDEDIYVNGKGYIYKDGKYYVYLKDA 140  
 |||||  
 Db 61 SHGDHYHYNGKPYDAIIESELLMKDPNKLKDEDIYVNGKGYIYKDGKYYVYLKDA 120  
 QY 141 AHADNVTKKEINROKQESHQREGGTPRNDGAVALARSGRTTDDGYTFENASDIIEPT 200  
 |||||  
 Db 121 AHADNVTKKEINROKQESHQREGGTPRNDGAVALARSGRTTDDGYTFENASDIIEPT 180  
 QY 201 GDAYIVPHGDHYHYIPKNELASASELAFAFLSGRNLNSRTYRRQNSDNTSRTNWVPS 260  
 |||||  
 Db 181 GDAYIVPHGDHYHYIPKNELASASELAFAFLSGRNLNSRTYRRQNSDNTSRTNWVPS 240  
 QY 261 VSNPCTTNTSNNSTNSQASQNDIDSLKQLYKPLSQHVESDGLVFPDQITTSRT 320  
 |||||  
 Db 241 VSNPCTTNTSNNSTNSQASQNDIDSLKQLYKPLSQHVESDGLVFPDQITTSRT 300  
 QY 321 ARGVAVPHGDHYHYIPYQSOMSELEERTARIITPLRYSNHNVPSRPEOPSPOPTPEPSR 380  
 |||||  
 Db 301 ARGVAVPHGDHYHYIPYQSOMSELEERTARIITPLRYSNHNVPSRPEOPSPOPTPEPSR 360  
 QY 381 POPAPNLKIDSNSLSQVLRKVGEGYVFEKGISRYVFAKDLPSFTVKNLESKLSKQES 440  
 |||||  
 Db 361 POPAPNLKIDSNSLSQVLRKVGEGYVFEKGISRYVFAKDLPSFTVKNLESKLSKQES 420  
 QY 441 VSHRTLTKKKNVAPRODEFDKNVNLTEAHKALFEKNGNSDQALDKLETLNDESN 500  
 |||||  
 Db 421 VSHRTLTKKKNVAPRODEFDKNVNLTEAHKALFEKNGNSDQALDKLETLNDESN 480  
 QY 501 KEKLVDDLFLAPITPHERLKPNSQIEYTEDEVRIQAQADKYTTSDGYIFDEHDIISD 560  
 |||||  
 Db 481 KEKLVDDLFLAPITPHERLKPNSQIEYTEDEVRIQAQADKYTTSDGYIFDEHDIISD 540  
 QY 561 EGDAYVTPHMGSHHWIGKDSLSDKEKVAQAQYTKKGIILPPSPDADYKANPTGDSAAAY 620  
 |||||  
 Db 541 EGDAYVTPHMGSHHWIGKDSLSDKEKVAQAQYTKKGIILPPSPDADYKANPTGDSAAAY 600  
 QY 621 NRVAEKRIPLVRLPYVEHTVEYKNGNLIIPKDHYNHKKFAMPDHTYKANGYTLED 680  
 |||||  
 Db 601 NRVAEKRIPLVRLPYVEHTVEYKNGNLIIPKDHYNHKKFAMPDHTYKANGYTLED 660  
 QY 681 LFATIKYVHPERPHSHNDGNGNASEHVLGKDKHSEDPKNKFADEEPEVETPAPEVP 740  
 |||||  
 Db 661 LFATIKYVHPERPHSHNDGNGNASEHVLGKDKHSEDPKNKFADEEPEVETPAPEVP 720  
 QY 741 QVELEKEAOLKEAEVLLAVYTSSSLKANATETLAGLRNNLTLOIMDNNSIMAEAKLLA 800  
 |||||  
 Db 721 QVELEKEAOLKEAEVLLAVYTSSSLKANATETLAGLRNNLTLOIMDNNSIMAEAKLLA 780  
 QY 801 LKGSNPSYSKERIN 816  
 |||||  
 Db 781 LKGSNPSYSKERIN 796

RESULT 2

US-09-884-465A-7  
 ; Sequence 7, Application US/09884465A  
 ; Publication No. US20030077293A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shire Biochem, Inc.  
 ; APPLICANT: Hamel, Josee  
 ; APPLICANT: Brodeur, Bernard  
 ; APPLICANT: Martin, Denis  
 ; APPLICANT: Charland, Nathalie

; APPLICANT: Ouellet, Catherine  
 ; TITLE OF INVENTION: Streptococcus Antigens  
 ; FILE REFERENCE: 055190-0044  
 ; CURRENT APPLICATION NUMBER: US/09/884,465A  
 ; CURRENT FILING DATE: 2001-06-20  
 ; PRIOR APPLICATION NUMBER: 60/212,683  
 ; PRIOR FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 384  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 7  
 ; LENGTH: 840  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus pneumoniae  
 US-09-884-465A-7

Query Match 77.5%; Score 3304; DB 9; Length 840;  
 Best Local Similarity 75.3%; Pred. No. 1.6e-200;  
 Matches 632; Conservative 67; Mismatches 106; Indels 34; Gaps 3;

QY 1 MKINKYLVGSAALILVCSYEELGLYQARTVKEENNVSVIDKQATQKTENLTPEVSK 60  
 |||||  
 Db 1 MKINKYLVGSAALILVCSYEELGLYQARTVKEENNVSVIDKQATQKTENLTPEVSK 60  
 QY 61 REGINAEOIYIKITDQGYVTSBGHYYHYNGKVPYDAIIESELLMKDPNKLKDEDIYNE 120  
 |||||  
 Db 61 REGINAEOIYIKITDQGYVTSBGHYYHYNGKVPYDAIIESELLMKDPNKLKDEDIYNE 120  
 QY 121 VKGGVYVWDGKYVYVYLDAAHADNVTKKEINROKQESHQREGGTPRNDGAVALARSQ 180  
 |||||  
 Db 121 VKGGVYVWDGKYVYVYLDAAHADNVTKKEINROKQESHQREGGTPRNDGAVALARSQ 180  
 QY 181 GRYYTDDGYIFENASDIIEPTGDAYIVPHGDHYHYIPKNELASASELAFAFLSGRNLNS 240  
 |||||  
 Db 181 GRYYTDDGYIFENASDIIEPTGDAYIVPHGDHYHYIPKNELASASELAFAFLSGRNLNS 240  
 QY 241 SRTYRRQNSDNTSRTNWVPSVSNPCTTNTSNNSTNSQASQNDIDSLKQLYKPLS 300  
 |||||  
 Db 241 SRTYRRQNSDNTSRTNWVPSVSNPCTTNTSNNSTNSQASQNDIDSLKQLYKPLS 300  
 QY 301 ORHVESDGLVFPDQITTSRTAGVAVPHGDHYHYIPYQSOMSELEERTARIITPLRYSNHW 360  
 |||||  
 Db 301 ORHVESDGLVFPDQITTSRTAGVAVPHGDHYHYIPYQSOMSELEERTARIITPLRYSNHW 360  
 QY 361 VPDSPPEPSPOPTPEPSPOPTPEPSPOPTPEPSPOPTPEPSPOPTPEPSPOPTPEPS 416  
 |||||  
 Db 361 VPDSPPEPSPOPTPEPSPOPTPEPSPOPTPEPSPOPTPEPSPOPTPEPSPOPTPEPS 416  
 QY 417 VYFANDLPSEIYKNEKLSKQESVSHLTAKKEVAVARDEFDKCANLLTEAHKALFE 476  
 |||||  
 Db 417 VYFANDLPSEIYKNEKLSKQESVSHLTAKKEVAVARDEFDKCANLLTEAHKALFE 476  
 QY 476 YLPARNLSAETAGIDSKLAKQESLSHKLGAKKTDLPSSDRPEFNKAYDILLARIHQDILLD 480  
 |||||  
 Db 476 YLPARNLSAETAGIDSKLAKQESLSHKLGAKKTDLPSSDRPEFNKAYDILLARIHQDILLD 480  
 QY 477 NGRNSDQALDKLEERLNDESTNKEKLVDDLFLAPITPHERLKPNSQIEYTEDEVRI 536  
 |||||  
 Db 477 NGRNSDQALDKLEERLNDESTNKEKLVDDLFLAPITPHERLKPNSQIEYTEDEVRI 536  
 QY 481 NKGROYDFEALDNLLEERLKDVSDDKVKLVDDILAFLAIRPHERLKPNSQIETDDEIQ 540  
 |||||  
 Db 481 NKGROYDFEALDNLLEERLKDVSDDKVKLVDDILAFLAIRPHERLKPNSQIETDDEIQ 540  
 QY 537 IQAQADKYTTSDGYIFDEHDIISDEGDAYVTPHMGSHHWIGKDSLSDKEKVAQAQYTKK 596  
 |||||  
 Db 537 IQAQADKYTTSDGYIFDEHDIISDEGDAYVTPHMGSHHWIGKDSLSDKEKVAQAQYTKK 596  
 QY 541 VAKLAGKYTEDEGYTFEDRDITSDDEGDAYVTPHMTSHHWIKKDSIEAEKAAQAAYAKK 600  
 |||||  
 Db 541 VAKLAGKYTEDEGYTFEDRDITSDDEGDAYVTPHMTSHHWIKKDSIEAEKAAQAAYAKK 600  
 QY 597 GLPPSPDADYKANPTGSAALYVYVNGKRIPLVRLPYVEHTVEYKNGNLIIPKDH 656  
 |||||  
 Db 597 GLPPSPDADYKANPTGSAALYVYVNGKRIPLVRLPYVEHTVEYKNGNLIIPKDH 656  
 QY 601 GLPPSPDADYKANPTGSAALYVYVNGKRIPLVRLPYVEHTVEYKNGNLIIPKDH 660  
 |||||  
 Db 601 GLPPSPDADYKANPTGSAALYVYVNGKRIPLVRLPYVEHTVEYKNGNLIIPKDH 660  
 QY 657 YNINIKFAFEDDHYKAPNGYTLEDLFATIKYVHPERPHSHNDGNGNASEHVLGKDKH 716  
 |||||  
 Db 657 YNINIKFAFEDDHYKAPNGYTLEDLFATIKYVHPERPHSHNDGNGNASEHVLGKDKH 716  
 QY 661 YNINIKFAFEDDHYKAPNGYTLEDLFATIKYVHPERPHSHNDGNGNASEHVLGKDKH 720  
 |||||  
 Db 661 YNINIKFAFEDDHYKAPNGYTLEDLFATIKYVHPERPHSHNDGNGNASEHVLGKDKH 720  
 QY 717 EDPNKNFKAD-----EEPEVETPAPEVPQVETEK 746  
 |||||  
 Db 717 EDPNKNFKAD-----EEPEVETPAPEVPQVETEK 746  
 QY 747 VEAQLEKEAEVLLAKYTSSSLKANATETLAGLRNNLTLOIMDNNSIMAEAKLLKGS 805  
 |||||

```

Db      781  VEKKREAEDLLGKIQDILIKSAKETILTGLKNNLLFGQDNNNTIMAEKRLALLKES 839

RESULT 3
US-09-884-465A-8
: Sequence 8, Application US/09884465A
: Publication No. US20030077293A1
: GENERAL INFORMATION:
:   APPLICANT: Shire Biochem, Inc.
:   APPLICANT: Hamel, Josee
:   APPLICANT: Brodeur, Bernard
:   APPLICANT: Martin, Denis
:   APPLICANT: Charland, Nathalie
:   APPLICANT: Ouellet, Catherine
:   TITLE OF INVENTION: Streptococcus Antigens
:   FILE REFERENCE: 05190-004
:   CURRENT APPLICATION NUMBER: US/09/884, 465A
:   CURRENT FILING DATE: 2001-06-20
:   PRIOR APPLICATION NUMBER: 60/212,683
:   PRIOR FILING DATE: 2000-06-20
:   NUMBER OF SEQ ID NOS: 384
:   SOFTWARE: PatentIn version 3.1
:   SEQ ID NO 8
:   LENGTH: 838
:   TYPE: PRT
:   ORGANISM: Streptococcus pneumoniae
US-09-884-465A-8

Query Match      66.4%; Score 2832.5; DB 9; Length 838;
Best Local Similarity 64.9%; Pred. No. 9.5e-17;
Matches 554; Conservative 91; Mismatches 148; Indels 61; Gaps

QY      1  MKINKYLVGSAALILVSCSYELGLYQARY-KENNRKSYIDGQAOYTKTNLPDEV 59
      1  MKINKYLVGSAVAALVSVCSYELGRHOAGQYKKSNNRSYIDGQAGAKANLTPDEV 60
QY      60  KREGINAQIYKIKIDGQYVTSHGDIHYHYNGKVPYDAIISELLMKDPNYKIKDEY 119
      61  KREGINAQIYKIKIDGQYVTSHGDIHYHYNGKVPYDAIISELLMKDPNYKIKDEY 120
QY      120  EYKGGYVIRKVDGKYVYLKDAAHADNVRTKEEINRQKQEHSGHREGGPRNDGAVALARS 179
      121  EIKGGYVIRKVDGKYVYLKDAAHADNVRTKEEIKRQKQEHSHNHS--RADNVAVAARA 177
QY      180  QGRYTTDGGYIFNADIIEDTGDAIVPHGDIHYHTPKNELSASELAAAEALFSGGNIS 239
      178  QGRYTTDGGYIFNADIIEDTGDAIVPHGDIHYHTPKNELSASELAAAEAWNG---- 232
QY      240  NSRTYRNSDNTSTNNVPSVSNPGTNTNTSNNSTNSQASQSDIDSLIKOLYKLP 299
      233  -----KQGSRPSSSSSTYNANVPQRLSENHNLVTPYTHQ--NQGBNISSLIRELYAKPL 285
QY      300  SQRHVESDGLVFDPQOITSRTARGAIVPHGDIHYHTPYQSOMSELEERTARIPLKRYSNH 359
      286  SERHVESDGLVFDPQOITSRTARGAIVPHGDIHYHTPYQSOMSELEERTARIPLKRYSNH 345
QY      360  WVPDSRPQPSQPPIPPPS-----PGQPAPNLKIDSNSLSVLSOLVRRVGGYVFEKG 413
      346  WVPDSRPQPSQPPIPPPS-----PGQPAPNLKIDSNSLSVLSOLVRRVGGYVFEENG 403
QY      414  ISRYFAFDIDSEYTKNLESKTSKQESVSHLTAKKENAPAPDQDFYKANYLLLEAHKA 473
      404  VSRYPADLSLSEITAGIDSKLAKESHSKLGAKRTDLPSSDREYENKAYLLLRITTD 463
QY      474  LFENKGRSDFOALDKLERLNDSESTNKEKLVDDDLAFLAPITPHEBRLGKPSQSLFYETD 533
      464  LLDNKGQGVDFEVDNLLERLKDVSDKAYKLVDDLLAFLAPITPHEBRLGKPSQSLFYETD 523
QY      534  EYRIAOLADKYTTSDGYIFDEHDIISDEGDAVYTPHMGSHWIKGDSLSDKREKVAQAAYT 593
      524  EIOVAKLAGKYTTEDGYIFDPRDITSDEGDAVYTPHMTSHWIKKDSISEARAAQAAYA 583
QY      594  KEKGLPPSPADVYKANTPGDSAAIITRYKAGEKRIPLVRLPYWVEHYEVANGNLIIPH 653

```

Db	564	KEKGLTPSTHDQDSGNTKAEKAEALTYNRVKAQKVPFLDRMYNLQTYEYVANGSSILIPH	643	
QY	654	KDHYHNKIFWAMEDDHTYKAPNGYTTLEDDLFATIKYVVEHPDERPHSNDGNGNSEHYLGR	712	
Db	644	YDHYHNKIFWEMDEGLYEAHPKGYSLIEDLLATKYVVEHPNERPHSNDGNGNSADHYRKAK	703	
QY	713	-----KHSE-----	-----DPNNKFKADEPVEETPAE	736
Db	704	ADQDSKPEDKCHDEVSSEPTHPESDEKENHAGLNPSSADNLVYKPSYDTEPTEFEADDTTDE	763	
QY	737	PEVPOVETEKYEAOLKEAEVILLAKYTDSLSKANKATETLAGLKNNTLTQIMDNNSIMAAE	796	
Db	764	AEIPQVENSVIAKIADAEALKEKVTDSIRONAMETTLGLKSSLLGTKDNNTISAEDV	823	
QY	797	KLALLKGSNPSSV 810		
Db	824	SLALLKESQPAFI 837		
RESULT 4				
US-09-765-272-66				
; Sequence 66, Application US/09765272				
; Patent No. US20020061545A1				
GENERAL INFORMATION:				
APPLICANT: Choi et. al.				
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines				
NUMBER OF SEQUENCES: 452				
CORRESPONDENCE ADDRESS:				
ADDRESSEE: Human Genome Sciences, Inc.				
STREET: 9410 Key West Avenue				
CITY: Rockville				
STATE: Maryland				
COUNTRY: USA				
ZIP: 20850				
COMPUTER READABLE FORM:				
MEDIUM TYPE: diskette, 3.50 inch, 1.4mb storage				
COMPUTER: HP Vectra 486/33				
OPERATING SYSTEM: MSDOS version 6.2				
SOFTWARE: ASCII Text				
CURRENT APPLICATION DATA:				
APPLICATION NUMBER: US/09/765, 272				
FILING DATE: 22-Jan-2001				
CLASSIFICATION: <unknown>				
PRIOR APPLICATION DATA:				
APPLICATION NUMBER: 08/961,083				
FILING DATE: <unknown>				
ATTORNEY/AGENT INFORMATION:				
NAME: Brookes, A. Anders				
REGISTRATION NUMBER: 36,373				
REFERENCE/DOCKET NUMBER: PB340P2				
TELECOMMUNICATION INFORMATION:				
TELEPHONE: (301) 309-8504				
TELEFAX: (301) 309-8512				
INFORMATION FOR SEQ ID NO: 66:				
SEQUENCE CHARACTERISTICS:				
LENGTH: 763 amino acids				
TYPE: amino acid				
STRANDEDNESS: single				
TOPOLOGY: linear				
MOLECULE TYPE: protein				
SEQUENCE DESCRIPTION: SEQ ID NO: 66:				
US-09-765-272-66				
Query Match				
Best Local Similarity 62.4%; Score 2661.5; DB 10; Length 763;				
Matches 520; Conservative 78; Mismatches 131; Indels 49; Gaps 8;				
QY	20	CSYELGLVQARIV-KENNRSVYTDCKQATQKQENTLPDEVSRRREGIMAEQIYIKITDQSY 78		
Db	1	CSYELGRHQAGQYKRESNRSYIIDDDQAGQKAKENLTPDEVSRRREGIMAEQYIKITDQSY 60		
QY	79	VTSHGDHYHYNGKPYPAIISELLMDPNYKDKDEDIVNEVGKGYIVIKVDGKYVYVVK 138		



```
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 382
; LENGTH: 1365
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: xaa = Methionine or nothing
; NAME/KEY: MISC_FEATURE
; LOCATION: (357)..(357)
; OTHER INFORMATION: xaa = Glycine or nothing
; NAME/KEY: MISC_FEATURE
; LOCATION: (558)..(558)
; OTHER INFORMATION: xaa = Proline or nothing
US-09-884-465A-382
```

```
Query Match          44.4%; Score 1894.5; DB 9; Length 1365;
Best Local Similarity 65.0%; Pred. No. 3.3e-111;
Matches 367; Conservative 61; Mismatches 98; Indels 39; Gaps 4;
```

```
QY 286 DIDSLLKQLYKLPESQHNVESDGLVDPQAQITSRTARGVAAPHGHYHFIPTYSQSELEE 345
D 2 DIDSLLKQLYKLPESQHNVESDGLVDPQAQITSRTARGVAAPHGHYHFIPTYSQSELEE 61
QY 346 RIARIIPLRKSNHWPDSRPEQSPQPTPEPSPQAPAPNLK-IDSN--SSLVSQLR 401
D 62 RIARIIPLRKSNHWPDSRPEQSPQPTPEPSPQAPAPNPAPNSNPIDEXLVEAYR 121
QY 402 KVGEGYVEEKGISRYVFAKDLPEFTVKNLESKLSKQESVSHTLAKKENVAPRQGEFYD 461
D 122 KVGEGYVEEKGISRYVFAKDLPEFTVKNLESKLSKQESVSHTLAKKENVAPRQGEFYD 181
QY 462 KAVNLTLEAKALFEKNGRNSDFOALDKLERLNDESTNKEKLVDDLAFAPITRPEL 521
D 182 KAVNLTLEAKALFEKNGRNSDFOALDKLERLNDESTNKEKLVDDLAFAPITRPEL 241
QY 522 GRKNSOLEYTEDEVRIQAOLADKYTTSQGYTFDEHDIISDEGDAYVTPHMGSHWIKD 581
D 242 GRKNSOLEYTEDEVRIQAOLADKYTTSQGYTFDEHDIISDEGDAYVTPHMGSHWIKD 301
QY 582 SDEKVAQAQYATTEKGIPLPSPDADYKANPTGSAATYNNRYKGEKRITLVLPYVHEHT 641
D 302 SDEKVAQAQYATTEKGIPLPSPDADYKANPTGSAATYNNRYKGEKRITLVLPYVHEHT 361
QY 642 VEKKNGLIIPHDHYNHNIKFAFDDHTYKAPNGYTLLEDFATIKYVEHPDERPHSNDG 701
D 362 VEKKNGLIIPHDHYNHNIKFAFDDHTYKAPNGYTLLEDFATIKYVEHPDERPHSNDG 421
QY 702 WGNASEHVILGKKDHSDEPNKFNKAD-----EEPE 731
D 422 WGNASEHVILGKKDHSDEPNKFNKAD-----EEPE 481
QY 732 ETDAPEPEVPOVETEKVEAQLKEAEVLLAKVTSSKANATETLAGLRNNLTQIMDNNSI 791
D 482 ETDAPEPEVPOVETEKVEAQLKEAEVLLAKVTSSKANATETLAGLRNNLTQIMDNNSI 541
QY 792 MAEAEKLLALLKSGS-----NPSVS 811
D 542 MAEAEKLLALLKSGS-----NPSVS 566
```

```
RESULT 7
US-09-884-465A-380
; Sequence 380, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
```

```
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0004
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 380
; LENGTH: 1139
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: xaa = Methionine or nothing
; NAME/KEY: MISC_FEATURE
; LOCATION: (570)..(570)
; OTHER INFORMATION: xaa = Glycine or nothing
; NAME/KEY: MISC_FEATURE
; LOCATION: (571)..(571)
; OTHER INFORMATION: xaa = Proline or nothing
US-09-884-465A-380
```

```
Query Match          42.2%; Score 1798.5; DB 9; Length 1139;
Best Local Similarity 60.7%; Pred. No. 3e-105;
Matches 345; Conservative 72; Mismatches 106; Indels 45; Gaps 4;
```

```
QY 286 DIDSLLKQLYKLPESQHNVESDGLVDPQAQITSRTARGVAAPHGHYHFIPTYSQSELEE 345
D 3 NISLLRELKALPLSERHVESDGLIFDPAQITSRTARGVAAPHGHYHFIPTYSQSELEE 62
QY 346 RIARIIPLRKSNHWPDSRPEQSPQPTPEPSPQAPAPNLK-IDSN--SSLVSQLR 399
D 63 RIARIIPLRKSNHWPDSRPEQSPQPTPEPSPQAPAPNPAPNSNPIDEXLVEAYR 120
QY 400 KVGEGYVEEKGISRYVFAKDLPEFTVKNLESKLSKQESVSHTLAKKENVAPRQGEFYD 459
D 121 KVGEGYVEEKGISRYVFAKDLPEFTVKNLESKLSKQESVSHTLAKKENVAPRQGEFYD 180
QY 460 YKAVNLTLEAKALFEKNGRNSDFOALDKLERLNDESTNKEKLVDDLAFAPITRPEL 519
D 181 YKAVNLTLEAKALFEKNGRNSDFOALDKLERLNDESTNKEKLVDDLAFAPITRPEL 240
QY 520 GRKNSOLEYTEDEVRIQAOLADKYTTSQGYTFDEHDIISDEGDAYVTPHMGSHWIKD 579
D 241 GRKNSOLEYTEDEVRIQAOLADKYTTSQGYTFDEHDIISDEGDAYVTPHMGSHWIKD 300
QY 580 SDEKVAQAQYATTEKGIPLPSPDADYKANPTGSAATYNNRYKGEKRITLVLPYVHEHT 639
D 301 SDEKVAQAQYATTEKGIPLPSPDADYKANPTGSAATYNNRYKGEKRITLVLPYVHEHT 360
QY 640 HIVEKKNGLIIPHDHYNHNIKFAFDDHTYKAPNGYTLLEDFATIKYVEHPDERPHSNDG 699
D 361 HIVEKKNGLIIPHDHYNHNIKFAFDDHTYKAPNGYTLLEDFATIKYVEHPDERPHSNDG 420
QY 700 DGMNASEHVILGKKDHSDEPNKFNKAD-----DPKN 722
D 421 DGMNASEHVILGKKDHSDEPNKFNKAD-----DPKN 480
QY 723 FKADEPEPEVPOVETEKVEAQLKEAEVLLAKVTSSKANATETLAGLRNNLT 782
D 481 FKADEPEPEVPOVETEKVEAQLKEAEVLLAKVTSSKANATETLAGLRNNLT 540
QY 783 LOIMDNNSIMAEKLLALLKSGSNPSVS 810
D 541 LOIMDNNSIMAEKLLALLKSGSNPSVS 568
```

RESULT 8

US-09-884-465A-378  
Sequence 378, Application US/09884465A  
Publication No. US20030077293A1  
GENERAL INFORMATION:  
APPLICANT: Shire Biochem, Inc.  
APPLICANT: Hamel, Josee  
APPLICANT: Brodeur, Bernard  
APPLICANT: Martin, Denis  
APPLICANT: Charland, Nathalie  
APPLICANT: Ouellet, Catherine  
TITLE OF INVENTION: Streptococcus Antigens  
FILE REFERENCE: 055190-0044  
CURRENT APPLICATION NUMBER: US/09/884,465A  
CURRENT FILING DATE: 2001-06-20  
PRIOR APPLICATION NUMBER: 60/212,683  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 384  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 378  
LENGTH: 1378  
TYPE: PRF  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Unknown Organism  
NAME/KEY: MISC\_FEATURE  
LOCATION: (1)..(1)  
OTHER INFORMATION: Xaa = Methionine or nothing  
NAME/KEY: MISC\_FEATURE  
LOCATION: (570)..(570)  
OTHER INFORMATION: Xaa = Glycine or nothing  
NAME/KEY: MISC\_FEATURE  
LOCATION: (571)..(571)  
OTHER INFORMATION: Xaa = Proline or nothing  
US-09-884-465A-378

Query Match 42.2%; Score 1798.5; DB 9; Length 1378;  
Best Local Similarity 60.7%; Pred. No. 3.9e-105;  
Matches 345; Conservative 72; Mismatches 106; Indels 45; Gaps 4;

QY 286 DIDSLKQLKLPISQKHVESDGLVDFDPAQITSTRAGVAVPHGDHYHFTFYQSMSLEE 345  
DB 3 NISLRELAKLPISQKHVESDGLVDFDPAQITSTRAGVAVPHGDHYHFTFYQSMSLEE 62  
QY 346 RIARILPRYSNMHWPPSRPEQSPQTPPEPS-----PGOPAPNLKIDSNSLVSQL 399  
DB 63 RIARILPRYSNMHWPPSRPEQSPQTPPEPSQOPAPNPQAPNSPID--EKLVEKA 120  
QY 400 VRAVGEVVEEKGISRYFAKDLPSSETVKNLESKSKOESVSHLTJAKKENVAPRDOEF 459  
DB 121 VRAVGEVVEEKGISRYFAKDLPSSETVKNLESKSKOESVSHLTJAKKENVAPRDOEF 180  
QY 460 YKRAYNLTEAHKALEENKGRNSDFQALDKLERLNDSTNKEKLVDDLLAFLAPITTHEP 519  
DB 181 YKRAYNLTEAHKALEENKGRNSDFQALDKLERLNDSTNKEKLVDDLLAFLAPITTHEP 240  
QY 520 RLGKPNQSIYEYDEVRIAQADKYTTSOGYTFDEHDIISDEGDAYVTPHMGSHWIGKD 579  
DB 241 RLGKPNQSIYEYDEVRIAQADKYTTSOGYTFDEHDIISDEGDAYVTPHMGSHWIGKD 300  
QY 580 SLSDKEVNAQAATYKKEGILPRSPDADYKANPTGDSAAALYNNVKEEKLPLVRLPYMVE 639  
DB 301 SLSDKEVNAQAATYKKEGILPRSPDADYKANPTGDSAAALYNNVKEEKLPLVRLPYMVE 360  
QY 640 HWEVKNKGNLIIPKDHYNHIFKAFMDHTYKAPNGYTTLEDLPATIKYVVEHDERPHSNG 699  
DB 361 HWEVKNKGNLIIPKDHYNHIFKAFMDHTYKAPNGYTTLEDLPATIKYVVEHDERPHSNG 420  
QY 700 DGMGNASEHVLGR-----KDHSE-----DPNKN 722  
DB 421 NGGNSNSDHYKRNKAKODSKPRDDEKHEDEYSEPTTHEPSEDEKENHAGLNSADNLKYPSPID 480  
QY 723 FKADPEVEETPAEPVPOVETEKVEAQLEAVLLAKYTTSDLSKANATETLAGLRNNLT 782  
DB 723 FKADPEVEETPAEPVPOVETEKVEAQLEAVLLAKYTTSDLSKANATETLAGLRNNLT 782

DB 481 TEETEAEADTTDEAEIPOVENSVINAKIADAEALLKVKTDPSIRONAMETTLGKSSLL 540  
QY 783 LOIMDNNSIMAEKELALLKGSNPSV 810  
DB 541 LGTKDNNTISAEBVDSLLALLKESQAPF 568

## RESULT 9

US-09-884-465A-381  
Sequence 381, Application US/09884465A  
Publication No. US20030077293A1  
GENERAL INFORMATION:  
APPLICANT: Shire Biochem, Inc.  
APPLICANT: Hamel, Josee  
APPLICANT: Brodeur, Bernard  
APPLICANT: Martin, Denis  
APPLICANT: Charland, Nathalie  
APPLICANT: Ouellet, Catherine  
TITLE OF INVENTION: Streptococcus Antigens  
FILE REFERENCE: 055190-0044  
CURRENT APPLICATION NUMBER: US/09/884,465A  
CURRENT FILING DATE: 2001-06-20  
PRIOR APPLICATION NUMBER: 60/212,683  
PRIOR FILING DATE: 2000-06-20  
NUMBER OF SEQ ID NOS: 384  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 381  
LENGTH: 1238  
TYPE: PRF  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Unknown Organism  
NAME/KEY: MISC\_FEATURE  
LOCATION: (1)..(1)  
OTHER INFORMATION: Xaa = Methionine or nothing  
NAME/KEY: MISC\_FEATURE  
LOCATION: (430)..(430)  
OTHER INFORMATION: Xaa = Glycine or nothing  
NAME/KEY: MISC\_FEATURE  
LOCATION: (431)..(431)  
OTHER INFORMATION: Xaa = Proline or nothing  
US-09-884-465A-381

Query Match 39.1%; Score 1668; DB 9; Length 1238;  
Best Local Similarity 66.5%; Pred. No. 5.8e-97;  
Matches 314; Conservative 58; Mismatches 96; Indels 4; Gaps 2;

QY 286 DIDSLKQLKLPISQKHVESDGLVDFDPAQITSTRAGVAVPHGDHYHFTFYQSMSLEE 345  
DB 2 DIDSLKQLKLPISQKHVESDGLVDFDPAQITSTRAGVAVPHGDHYHFTFYQSMSLEE 61  
QY 346 RIARILPRYSNMHWPPSRPEQSPQTPPEPSQOPAPNLK--IDSN--SLVSQLVR 401  
DB 62 RIARILPRYSNMHWPPSRPEQSPQTPPEPSQOPAPNPQAPNSPID--EKLVEKA 121  
QY 402 KYGEGVVEEKGISRYFAKDLPSSETVKNLESKSKOESVSHLTJAKKENVAPRDOEF 461  
DB 62 KYGEGVVEEKGISRYFAKDLPSSETVKNLESKSKOESVSHLTJAKKENVAPRDOEF 121  
QY 462 KAYNLTEAHKALEENKGRNSDFQALDKLERLNDSTNKEKLVDDLLAFLAPITTHEP 521  
DB 182 KAYNLTEAHKALEENKGRNSDFQALDKLERLNDSTNKEKLVDDLLAFLAPITTHEP 241  
QY 522 GKPNSQIYEYDEVRIAQADKYTTSOGYTFDEHDIISDEGDAYVTPHMGSHWIGKD 581  
DB 242 GKPNSQIYEYDEVRIAQADKYTTSOGYTFDEHDIISDEGDAYVTPHMGSHWIGKD 301  
QY 582 SDKEVNAQAATYKKEGILPRSPDADYKANPTGDSAAALYNNVKEEKLPLVRLPYMVE 641  
DB 302 SDKEVNAQAATYKKEGILPRSPDADYKANPTGDSAAALYNNVKEEKLPLVRLPYMVE 361  
QY 642 VEKKNKGNLIIPKDHYNHIFKAFMDHTYKAPNGYTTLEDLPATIKYVVEHDERPHSNDG 701  
DB 642 VEKKNKGNLIIPKDHYNHIFKAFMDHTYKAPNGYTTLEDLPATIKYVVEHDERPHSNDG 701



```
Db 362 VEKNGSLIIPHDYHNIFEMFDEGLYEAPKGYTLEDLATATVKYVEHPNERPHSDNG 421
QY 702 WGNASEHVLGKKDSEDPNKNFKADEEPEVEETPAEPEVEKVEAOKE 753
Db 422 FGNASDHYXXNMOPQSLSYSTASDNTQSVAKGISTKPAKNSENLQSLKE 473

RESULT 10
US-09-884-465A-376
; Sequence 376, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 376
; LENGTH: 999
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Unknown Organism
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Methionine or nothing
; NAME/KEY: MISC_FEATURE
; LOCATION: (430)..(430)
; OTHER INFORMATION: Xaa = Glycine or nothing
; NAME/KEY: MISC_FEATURE
; LOCATION: (431)..(431)
; OTHER INFORMATION: Xaa = Proline or nothing
US-09-884-465A-376

Query Match 38.9%; Score 1659; DB 9; Length 999;
Best Local Similarity 61.9%; Pred. No. 1,66-96;
Matches 326; Conservative 63; Mismatches 98; Indels 40; Gaps 8;

QY 286 DIDSLLKQLYKPLSQRHVESDGLVFPDPAQITSRARGVAVPHGHHYFIPYSOMSELE 345
Db 2 DIDSLLKQLYKPLSQRHVESDGLVFPDPAQITSRARGVAVPHGHHYFIPYSOMSELE 61
QY 346 RIARIILPRYSNMHWPPDSRPEQSPQTPPEPSPGPAPNLK-IDSN--SSLVSQLV 401
Db 62 RIARIILPRYSNMHWPPDSRPEQSPQTPPEPSPGPAPNLK-IDSN--SSLVSQLV 121
QY 402 KVGEGYVEEKGISRYFAKDLPSFTVKNLESKLSQESVSHTLTAKKENAVAPRDQEFYD 461
Db 122 KVGEGYVEEKGISRYFAKDLPSFTVKNLESKLSQESVSHTLTAKKENAVAPRDQEFYD 181
QY 462 KAYNLITFAKALFENKGRNSDFQALDKLLERLNDESTNKEKLVDDLLAFAPITHPERL 521
Db 182 KAYNLITFAKALFENKGRNSDFQALDKLLERLNDESTNKEKLVDDLLAFAPITHPERL 241
QY 522 GKPNISOIYTEDEVRIAGLADKYTTSOGYTFDEHDIISDECDAYVTPHMGSHWIGKDSL 581
Db 242 GKPNISOIYTEDEVRIAGLADKYTTSOGYTFDEHDIISDECDAYVTPHMGSHWIGKDSL 301
QY 582 SDKEKVAQAQYTKRGILPSPDADYKANPTGDSAAATYNNKGEKRIPLVRLPYVWEHT 641
Db 302 SDKEKVAQAQYTKRGILPSPDADYKANPTGDSAAATYNNKGEKRIPLVRLPYVWEHT 361
QY 642 VEKNGSLIIPHDYHNIFEMFDEGLYEAPKGYTLEDLATATVKYVEHPNERPHSDNG 701
Db 362 VEKNGSLIIPHDYHNIFEMFDEGLYEAPKGYTLEDLATATVKYVEHPNERPHSDNG 421
```

```
Db 362 VEKNGSLIIPHDYHNIFEMFDEGLYEAPKGYTLEDLATATVKYVEHPNERPHSDNG 421
QY 702 WGNASEHVLGKKDSEDPNKNFKADEEPEVEETPAEPEVEKVEAOKE 753
Db 422 FGNASDHY--XXDLTEE--QIKAAQKHLEFVKTSHNGDLDSISHEQYF-----GN 468

QY 751 LKAEVLLAKYDSSILKANATETTLAQLRNNTLTQIMDNNSIMAEAK 797
Db 468 AKEMDLDRKI-----EKKIAGIMKQYGVK---RESIVNKEK 503

RESULT 11
US-09-884-465A-377
; Sequence 377, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/212,683
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 377
; LENGTH: 999
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Unknown Organism
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Methionine or nothing
; NAME/KEY: MISC_FEATURE
; LOCATION: (570)..(570)
; OTHER INFORMATION: Xaa = Glycine or nothing
; NAME/KEY: MISC_FEATURE
; LOCATION: (571)..(571)
; OTHER INFORMATION: Xaa = Proline or nothing
US-09-884-465A-377

Query Match 38.7%; Score 1652; DB 9; Length 999;
Best Local Similarity 71.7%; Pred. No. 4,46-96;
Matches 307; Conservative 50; Mismatches 67; Indels 4; Gaps 2;

QY 286 DIDSLLKQLYKPLSQRHVESDGLVFPDPAQITSRARGVAVPHGHHYFIPYSOMSELE 345
Db 572 DIDSLLKQLYKPLSQRHVESDGLVFPDPAQITSRARGVAVPHGHHYFIPYSOMSELE 631
QY 346 RIARIILPRYSNMHWPPDSRPEQSPQTPPEPSPGPAPNLK-IDSN--SSLVSQLV 401
Db 632 RIARIILPRYSNMHWPPDSRPEQSPQTPPEPSPGPAPNLK-IDSN--SSLVSQLV 691
QY 402 KVGEGYVEEKGISRYFAKDLPSFTVKNLESKLSQESVSHTLTAKKENAVAPRDQEFYD 461
Db 692 KVGEGYVEEKGISRYFAKDLPSFTVKNLESKLSQESVSHTLTAKKENAVAPRDQEFYD 751
QY 462 KAYNLITFAKALFENKGRNSDFQALDKLLERLNDESTNKEKLVDDLLAFAPITHPERL 521
Db 752 KAYNLITFAKALFENKGRNSDFQALDKLLERLNDESTNKEKLVDDLLAFAPITHPERL 811
QY 522 GKPNISOIYTEDEVRIAGLADKYTTSOGYTFDEHDIISDECDAYVTPHMGSHWIGKDSL 581
Db 812 GKPNISOIYTEDEVRIAGLADKYTTSOGYTFDEHDIISDECDAYVTPHMGSHWIGKDSL 871
QY 582 SDKEKVAQAQYTKRGILPSPDADYKANPTGDSAAATYNNKGEKRIPLVRLPYVWEHT 641
Db 642 SDKEKVAQAQYTKRGILPSPDADYKANPTGDSAAATYNNKGEKRIPLVRLPYVWEHT 701
```

Db 872 SEAEAAOAYAKKGLTPPSTDHQDSCTEAKGAFAIYNRYKAKKVPFLDMRYNLOXT 931  
 Oy 642 VEKKNGLIIPKHDYHNITKFAWFDHTYKAPNGYTLLEDLFAITKYVEHPDERPSNDG 701  
 Db 932 VEKKNGLIIPHDYHNITKFEWFDGLYEAAPKGYTLLEDLFAITKYVEHPDERPSNDG 991  
 Oy 702 WGNASEHV 709  
 Db 992 FGNASDHY 999

RESULT 12

US-09-884-465A-6  
 : Sequence 6, Application US/09884465A  
 : Publication No. US20030077293A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Shire Biochem, Inc.  
 : APPLICANT: Hamel, Josee  
 : APPLICANT: Brodeur, Bernard  
 : APPLICANT: Martin, Denis  
 : APPLICANT: Charland, Nathalie  
 : APPLICANT: Ouellet, Catherine  
 : TITLE OF INVENTION: Streptococcus Antigens  
 : FILE REFERENCE: 055190-0044  
 : CURRENT APPLICATION NUMBER: US/09/884,465A  
 : PRIOR FILING DATE: 2001-06-20  
 : PRIOR APPLICATION NUMBER: 60/212,683  
 : NUMBER OF SEQ ID NOS: 384  
 : SOFTWARE: PatentIn version 3.1  
 : SEQ ID NO 6  
 : LENGTH: 1039  
 : TYPE: PRN  
 : ORGANISM: Streptococcus pneumoniae  
 : US-09-884-465A-6

Query Match 30.4%; Score 1295.5; DB 9; Length 1039;  
 Best Local Similarity 41.5%; Pred. No. 1.5e-73;  
 Matches 309; Conservative 99; Mismatches 185; Indels 151; Gaps 20;

Oy 1 MKINKYU-VGSAALISVSYELGLQATVKNRVSITDCKOKTOKTENTPDEVS 59  
 Db 1 MKFSKYIAAGSAIVSLICAYALNHRSEKDNKRVSVDSOSSOSENLTPDOVS 60  
 Oy 60 KREGINAOIYIKITDQGYVSHDGHYHYNGKVPYDAIISELLMKDPNYKLEDEDIVN 119  
 Db 61 OREGIOAOIYIKITDQGYVSHDGHYHYNGKVPYDAIISELLMKDPNYKLEDEDIVN 120  
 Oy 120 EVKGGYIVKDGKYYVYLKDAAHADNVRTKEINRKOESQHREGGPRNDGVALARS 179  
 Db 121 EVKGGYIVKDGKYYVYLKDAAHADNVRTKEINRKOESQHREGGPRNDGVALARS 176  
 Oy 180 OGRYTTDGGYIFNMSDITIEDGDAIVPHGHHYIPKNETSASELAALAEFLSGRGLS 239  
 Db 177 OGRYTTDGGYIFNMSDITIEDGDAIVPHGHHYIPKNETSASELAALAEFLSGRGLS 236  
 Oy 240 NSRTYRONSDNTSRFTNVPVSNNGTNTNTNNSNTNSOASNDIDSLKOLYKLP 299  
 Db 237 SOLSTYSSTRASN-----NTDSVAKGSTRKPNKSENLOSLKELTSPS 280  
 Oy 300 SORHVESDGLVDPQAQTSRTARCAVPHGHYHPIPSOMSELEFARIILPLRYSNH 359  
 Db 281 AQRYSSEGLVDPQAQTSRTARCAVPHGHYHPIPSOMSELEFARIILPLRYSNH 353  
 Oy 360 WVPDRPDPSPQPTPEPSPGPAPNLKIDSNSLVQLRVKGEVGFEEKISRYP 419  
 Db 334 -----ISGTGSTVSTNAK----- 346  
 Oy 420 AKDLPSETVKNLESLKSOESVSHLTAKKENVAPRDOEFYDKAVNLLTEAHKALFENKG 479  
 Db 347 -----PNEVVSISGLSSNNS-----LTSKELSSASDGYTFNPK-DIVEETATAVIVRHG 397  
 Oy 480 RNSDFOALDKLERLNDESTNKEKLYVDLLAFITPRLERKGNOSQIETEDEVRIAQ 539

Db 398 --DHPHYLPK-----SNQIGOPTLPNNSLATPSP-SLPINPGTSHEKHE----- 438  
 Oy 540 LADKTYTSDGYIFPEHNDIISDEGAYVTPHGHSHWIGKSLSPKEXVAQAATYKREKIL 599  
 Db 439 -----EDGFGFANRIIADESGFVMSHGDHNNHFFPKDLTEQIKAAOKHLE----- 467  
 Oy 600 PPSPDADYKAPPTG-DSAAAIYNRYK-----EKRIPLVRLPYV--EHTVEV 644  
 Db 488 -----VKTSHNGLDLSLSHEDYPCNAKEMKDLCKIEKIGIMKOYGVKRESIYN 540  
 Oy 645 KNGNLIIPKHDYHNITKFAWFDHTYKAPNGYTLLEDLFAITKYVEHPDERPSNDG-- 701  
 Db 541 KKNKNIIPHDYHNITKFEWFDGLYEAAPKGYTLLEDLFAITKYVEHPDERPSNDG-- 590  
 Oy 702 --WGNASEHVLCCKHSDPKNKF 723  
 Db 591 VYTGEEELNVVALLKNSSTFNQNF 614

RESULT 13

US-09-765-272-182  
 : Sequence 182, Application US/09765272  
 : Patent No. US20020061545A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Choi et. al.  
 : TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
 : NUMBER OF SEQUENCES: 452  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESS: Human Genome Sciences, Inc.  
 : STREET: 9410 Key West Avenue  
 : CITY: Rockville  
 : STATE: Maryland  
 : COUNTRY: USA  
 : ZIP: 20850

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage  
 COMPUTER: HP Vectra 486/33  
 OPERATING SYSTEM: MSDOS version 6.2  
 SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/765,272  
 FILING DATE: 22-Jan-2001  
 CLASSIFICATION: <unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/961,083  
 FILING DATE: <unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brookes, A. Anders  
 REGISTRATION NUMBER: 36,373  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (301) 309-8504  
 TELEFAX: (301) 309-8512  
 INFORMATION FOR SEQ ID NO: 182:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 447 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 182:  
 US-09-765-272-182  
 Query Match 27.8%; Score 1186.5; DB 10; Length 447;  
 Best Local Similarity 41.0%; Pred. No. 3.5e-67;  
 Matches 263; Conservative 62; Mismatches 110; Indels 207; Gaps 11;

```

Db 61 DHYHYNGKYDYDALFSEELLMDPNYQOLKADIVNEVKGYIIVKDGKYYVIAKDAHA 120
QY 144 DNVPTKEINROKHOHSGHGGTFRNDGAVALARSGRTTDDGYIFPNSADIETDGA 203
Db 121 DNVPTKEINROKHOHSGHGGTFRNDGAVALARSGRTTDDGYIFPNSADIETDGA 176
QY 204 YIVPGDHYHYIPKNEISASLAAAEAFSLGKGLNSRTYRRONSNTSRTMVPVSXN 263
Db 177 YIVPGDHYHYIPKNEISASLAAAEAFSLGKGLNSRTYRRONSNTSRTMVPVSXN 224
QY 264 PGTNTNTSNNSTNSQASQNDISLKLQYKLPISQHVESDGLVFPDQITTSRTARG 323
Db 225 -----NTOSVAKGSTSKPANKSENLOSILKELDYDSPAQRYSFESDGLVFPDQITTSRTARG 280
QY 324 VAVPHGDHYHYIPKNEISASLAAAEAFSLGKGLNSRTYRRONSNTSRTMVPVSXN 383
Db 281 VAVPHGDHYHYIPKNEISASLAAAEAFSLGKGLNSRTYRRONSNTSRTMVPVSXN 309
QY 384 APNLKIDSNSLVSQYKRVGEGYVEEKGISRYVPAKDLPSFTVNLKESKLSKOSVSH 443
Db 310 -----ISGTGSTVSTNAK-----PNEVSGLSGSSNPS----- 339
QY 444 TLTAKEKNVAPRQEFYDKAYNLTFAHKAFLFNKGRNSFOALDKLRLNDESTNKEK 503
Db 340 -LITTSK-----LSSA----- 349
QY 504 LVDDLAFIAPITHPERLGRKPSQIEYTEDVIRIAQLADKYTTSDDGYIFDEHDIISDEGD 563
Db 350 -----SDGYIINPRDIVEETAT-----TGDSA 616
QY 564 AYTTPHMGSHMIGKSLDSKEVAAQATKKEGILPSPDADVKANP-----TGDSA 616
Db 367 AYTTPHMGSHMIGKSLDSKEVAAQATKKEGILPSPDADVKANP-----TGDSA 616
QY 617 AATYRNVKGEKRIPLVRLPYMVEHTEVANGNLIIPKOHYH 658
Db 419 GFDANRIAE-----DESGFVNSHGDNH 442

```

```

; OTHER INFORMATION: Xaa = Proline or nothing
US-09-884-465A-384
Query Match
Best Local Similarity 55.4%; Pred. No. 2e-54; Length 913;
Matches 189; Conservative 45; Mismatches 70; Indels 37; Gaps 2;

QY 507 DILAFLAPITHPERLGRKPSQIEYTEDVIRIAQLADKYTTSDDGYIFDEHDIISDEGDYAV 566
Db 2 DILAFLAPITHPERLGRKPSQIEYTEDVIRIAQLADKYTTSDDGYIFDEHDIISDEGDYAV 61
QY 567 TTPHMGSHMIGKSLDSKEVAAQATKKEGILPSPDADVKANP-----TGDSA 626
Db 62 TTPHMGSHMIGKSLDSKEVAAQATKKEGILPSPDADVKANP-----TGDSA 626
QY 627 KRIPLVRLPYMVEHTEVANGNLIIPKDHYNIKFAFMDHDTTYKAPNGYTFEDIPATIK 686
Db 122 KRIPLVRLPYMVEHTEVANGNLIIPKDHYNIKFAFMDHDTTYKAPNGYTFEDIPATIK 181
QY 687 YVEHPDERPHSDNGFVNAADHVRKAKADODSKPDEDEKHEDEVSPTHPESDEKENHAGL 717
Db 182 YVEHPDERPHSDNGFVNAADHVRKAKADODSKPDEDEKHEDEVSPTHPESDEKENHAGL 717
QY 718 -----DPNKNFKADEEPEVETPAPEPEVQVENEKEVQAEVLAATVTSLSKAN 769
Db 242 NPSADNLKPKSTDTTETETEDTDEAFIPQVENSVINAKIADAEALLERKVTDPISIRON 301
QY 770 ATETLAGLRNNITLQIMDNSTIMAEKLLALGKSNPSV 810
Db 302 AMETTLGLKSSILLGTCKDNNTISAEVDSLALLKESQAPI 342

RESULT 15
US-09-884-465A-379
; Sequence 379, Application US/09884465A
; Publication No. US20030077293A1
; GENERAL INFORMATION:
; APPLICANT: Shire Biochem, Inc.
; APPLICANT: Hamel, Josee
; APPLICANT: Brodeur, Bernard
; APPLICANT: Martin, Denis
; APPLICANT: Charland, Nathalie
; APPLICANT: Ouellet, Catherine
; TITLE OF INVENTION: Streptococcus Antigens
; FILE REFERENCE: 055190-0044
; CURRENT APPLICATION NUMBER: US/09/884,465A
; PRIOR FILING DATE: 2001-06-20
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 379
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Unknown Organism
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Xaa = Methionine or nothing
; NAME/KEY: MISC_FEATURE
; LOCATION: (344)..(344)
; OTHER INFORMATION: Xaa = glycine or nothing
; NAME/KEY: MISC_FEATURE
; LOCATION: (345)..(345)
; OTHER INFORMATION: Xaa = Proline or nothing
US-09-884-465A-379

Query Match
Best Local Similarity 55.4%; Pred. No. 2.8e-54; Length 1152;
Matches 189; Conservative 45; Mismatches 70; Indels 37; Gaps 2;

QY 507 DILAFLAPITHPERLGRKPSQIEYTEDVIRIAQLADKYTTSDDGYIFDEHDIISDEGDYAV 566

```

```

Db      2 DILAFLAPIRHPERLGGPNAQIITYTDDIOVAKLAGYTTEDGYLPDPDITSDSGDAY 61
QY      567 TPRMGSHWIGKDSLSDKEKVAQAAYTKKGLPPSPDADVKANPTGDSAAAIYNRKGE 626
Db      62 TPRMTHSHWIKKDSLSEAEKAAQAAYAKKGLPPSTDHODSGNTIENKGAFAIYNRKAA 121
QY      627 KRIPLVRLPYMEHTEVEKNGNLIIPHKHYNHIFAMFDDHTYKAPNGYTTLEDLPATIK 686
Db      122 KVPBLDRMPYNLOYTVEVKNKSLIIPYDHYNHIFEFWFDEGLYEAPKGYSLLEDLATYK 181
QY      687 YYVEHDERPHSNDGNGNASEHYLCK-----KDSE----- 717
Db      182 YYVEHDERPHSNDGNGNASEHYLCK-----KDSE----- 717
QY      718 -----DENKNFKADEEVEEETPAPEVQYETEKEVENQLEAEVLLAKYTDSLSKAN 769
Db      242 NPSADNLYKPSSTDEETEEDAEEDTDEAEIROYENSVINAKIADAEALLEKYTPDISRON 301
QY      770 ATETLAGLRNNLTQIMDNNSIMAEKLLALLKGSNPSV 810
Db      302 AMETLTGLKSSLLGTKNNTISAEVDSLALLKESOPAPI 342

```

Search completed: May 19, 2003, 08:24:12  
 Job time : 242.436 secs